

IEEE Signal Processing MAGAZINE

2011 年 IEEE 訊號處理學會論文摘錄
中文版



續約您的 IEEE 2011年會員 資格與加入訊號處理學會

訊號處理學會 (SPS) 會員權益有:

- ✓ 訂閱 IEEE 訊號處理雜誌
- ✓ 線上獲取所有 7 份 SPS 出版物
- ✓ 內容簡報，包括訊號處理出版物之目錄
- ✓ 內部訊號處理電子報(eNews)
- ✓ 學會會議註冊費的折扣
- ✓ 取得會議旅行補助金之資格



經濟套裝

US\$20/\$10 學生會員

包括：

- 訊號處理雜誌之數位遞送，每年6期
- 內部訊號處理電子報，每年12期

進階套裝

US\$35/\$18 學生會員

包括：

所有一級會員的福利外，

還有

- 訊號處理雜誌之列印版，每年6期
- 線上獲取 7 份 SPS 出版物
- 內容簡報之列印版，每年6期

當您今年秋天續約您的會員資格時，
請造訪：www.ieee.org/renew





訊號處理擴展至亞洲地區

我

謹代表IEEE
訊號處理學
會，藉此機
會向我們講
中文的同
事，學生和

更為廣泛的訊號處理雜誌讀者群致意。訊號處理是一個充滿活力的研究領域，在全球範圍內皆有研究和技術開發，特別是環太平洋地區。學會的出版物和會議都因來自此地地區的投稿而極大地豐富，透過建立分會，在此地區舉行研討會和會議，和其他志願活動而達到促進中文訊號處理社群的戰略目標。

我們中的很多人有幸參加 2010 年 9 月在香港舉辦的兩個重要會議中的一個。我由衷讚賞大會主席，蕭允治教授和他的同事們之進取精神，為這樣一個龐大的會議，在科學和文化兩方面所作出的巨大努力。無需說 2009 年在開羅，與 2010 年在香港舉辦的連續兩屆 ICIP 會議凸顯了訊號處理作為資訊時代的核心學科和技術，具有全球規模的生命力。更有在 ICASSP2009 之後很短時間內，在臺北進行的亞洲回訪中強調了學會對這些個活動，以及在此地區對這個領域所作巨大貢獻的認可。

誠如我在 2010 年 7 月的雜誌專欄中所述，學會正加倍努力，向分會和其他成員伸出援手，特別是在工業方面。例如，會議的副理事長，John

Mathews，在 2010 年夏季訪問印度的多個分會。同樣，在 ICIP2010 舉行的前一周，我們學會的數位志願領導者探訪中國大陸，與學生和專業人士會面，以鼓勵在該國建立更多的分會。

當選主席劉國瑞教授訪問廣州，並會見了來自中國大陸南部和西南部的其他有關各方。副財務長，吳旻教授也在廣州加入其中。廣州之行建立了一個新的，隸屬於北京 IEEE 分會，由黃繼武教授主持的訊號處理學會分會。期間，吳旻教授和我造訪北京，會見北京分會的主席，袁保宗教授，以及來自 IEEE 北京分會和中國電子學會（CIE）的其他領導人。9 月 21-22 日，我亦有拜訪由黃建國教授主持的西安分會。9 月 23-24 日，訊號處理雜誌的主編，鄧立，與我一起訪問上海，慶祝劉興釗教授主持的新的訊號處理學會分會的落成典禮。在此次訪問期間，兩位學會的志願者，周國彤教授和馬曉麗教授隨行左右。此行還探訪了北京的清華大學，中國科技大學（USTC），西安的西北工業大學，上海交通大學之師生。此次訪問以與訊號處理學會執行委員會和香港訊號處理學會分會的幾名會員的會談告一段落。

於我個人而言，這是一次順利訪問，在 1986 年中科院邀請聲學，語音和訊號處理學會約二十四年之後才能夠達成，在東京舉辦 ICASSP96 後，訪問上海和北京，並分別舉行訊

號處理會議。我們受到熱情款待，與中國同事交流資訊，並在 2010 年獲益良多。我們非常期待著繼續努力加強訊號處理學會的存在和知名度，增加來自亞洲 SPS 志願者的活躍會員並增進參與度。

【SP】



Enhancing the Global Impact of Our Magazine— Introducing the second reprint edition of SPM in Chinese

Last September, embracing IEEE's global strategy, IEEE Signal Processing Magazine (SPM) produced the first edition of translated reprint articles from SPM in simplified Chinese. It has been hosted on our society's website, available digitally [1], and its hard copies have been distributed to ICIP attendees in Hong Kong, September 26-29, 2010 with a formal follow-up survey. The feedbacks are strongly positive, with numerous suggestions for improving future editions. One of the feedbacks is that many readers of our SPM in Taiwan and Hong Kong do not read simplified Chinese. This prompts us to use traditional Chinese, occasionally mixed with simplified Chinese, in our second edition that you are reading now.

In addition to the formal survey, we also conducted a number of informal surveys via discussions with the Chinese readers of SPM. Many native Chinese readers suggested the use of side-by-side English and translated Chinese articles. They told us that since their past writing of technical signal processing papers has typically been in Chinese, side-by-side comparisons between the two languages will help them improve their English writing skills. We have included two such articles in this edition. In the future, we plan to include more side-by-side articles and more diverse topics.

In our introduction to the first edition of the Chinese translated reprint articles from SPM [2], we outlined

potential global impact of our translation effort. A few months since then, the significance has become even more apparent. China produces the largest engineering force in the world, with 1,003 universities providing engineering degrees, 3.7 million undergraduate engineering students, and 470,000 post graduate engineering students. IEEE's has experienced phenomenal growth in China in the last 4 years, with its membership increasing by 83% and the number of articles published increasing by 511%. The number of IEEE sponsored conferences has increased by 265%. The largest number of downloads from IEEE Xplore are from Asia (59%), followed by North America and Europe. Despite the popularity of English in China, language barriers still exist. One very interesting finding from our survey highlights the importance of our translation effort in an enlightening way. That is, many signal processing students and professionals in China find our translated articles in Chinese useful not because they do not read English well, but because they desire to write English better. Their career development requires publications of their research work in English. Having both English and their native Chinese versions of our SPM articles readily available to them would help them write signal processing articles in English better with similar styles.

Our translations capture cutting edge industry trends as well as tutorial articles in popular signal processing theory,

methods, and applications. In this second edition, we include the smart grid application and three carefully tutorial articles. To make the communication more effective, we make use of strategic partners from industry and academia. This translated edition in your hand has been mailed to all IEEE members in mainland China, Taiwan, Hong Kong and Singapore along with Spectrum magazine.

Our translation work has been guided by the general principle of "Glocal = Global + Local", where a global mind set is customized to local taste and culture. While you are reading and using our reprinted editions, please keep in mind this principle and, whenever possible, provide us with your feedback on how you would like the future editions to be so that they can help you.

REFERENCES:

- [1] [http://www.signalprocessing.org/uploads/docs/Chinese TranslationSPmagazine.pdf](http://www.signalprocessing.org/uploads/docs/Chinese%20TranslationSPmagazine.pdf)
- [2] <http://signalprocessingsociety.org/newsletter/introducing-a-translated-reprint-edition-of-ieee-signal-processing-magazine-2>

SP



擴大我們雜誌的全球影響力

一 有關訊號處理雜誌中文版之第二次重印說明

去

年九月，適逢 IEEE 之全球戰略機遇，我們的 IEEE 訊號處理雜誌 (SPM)

出版了本雜誌第一版簡體中文的翻譯重印文章。在我們學會的網站上即可找到這些文章的數位版 [1]，其複本已於 2010 年 9 月 26-29 日在香港分發與 ICIP 的與會者，其後附有一份正式的跟蹤調查表。意見回應非常積極，對提升今後的版本提出多項建議。其中一項反饋是很多訊號處理雜誌的讀者來自台灣和香港，無法閱讀簡體中文。這促使我們在第二版中使用正體，正如您現在正在閱讀的文章。

除了正式的調查問卷外，我們還透過與 SPM 在中國讀者進行討論，進行非正式的調查。很多土生土長的中國讀者建議使用文章的中文翻譯和英文並排版式。他們告訴我們說，因為他們過去總是使用中文撰寫專門性的訊號處理論文，兩種語言間的對照有助於他們提高英語寫作技能。在這一版中，我們把兩篇文章按照這一版式進行排列。在未來，如果對當前這些文章的回應很積極，我們將計劃加入更

多這樣的文章，同時也有更多不同類型的文章加入。

在第一版 SPM 中文翻譯摘錄文章 [2] 的前言中，我們描畫了我們的翻譯工作所具有的潛在全球影響。自那以後數月，其意義也愈加明晰可見。中國有全世界最大的工程技術隊伍，有 1003 所大學提供工程學位，370 萬工程專業的大學生和 47 萬工程專業的研究生。IEEE 在中國經歷了驚人的增長，過去四年間，其成員增加 83%，發表文章數量增加 511%。IEEE 主辦的會議增加 265%。從 IEEEExplore 下載的人數位居亞洲之首 (59%)，其次是北美和歐洲。儘管英語在中國很普及，但語言障礙依然存在。我們的調查問卷中有一個非常有趣的發現，以一種頗具啟發的方式凸顯了我們翻譯工作的重要性。也就是說，許多中國的從事訊號處理的學生和專業人士發現我們的翻譯文章很有用，並不是因為他們無法流暢地閱讀英語，而是因為他們希望能更好地撰寫英文。他們的職業發展需要用英語發表自己的研究工作。SPM 文章同時擁有英語和他們母語兩個版本，隨時可幫助他們更好地使用英語撰寫訊號處理方面類似風格的文章。

我們所翻譯的文章捕捉到最前沿的行業趨勢，以及熱門的訊號處理理論、方法和應用之指南文章。第二版中包括智慧電網應用和四篇精心挑選的指南性文章。為了更加有效地溝通，我們善加利用來自工業界和學術界的戰略合作夥伴。您手中的翻譯版已與頻譜雜誌一起遞送給所有在中國大陸，台灣，香港，新加坡的 IEEE 會員。

我們的翻譯工作一直遵循「全球在地化=全球+本地」的一般原則，其中全球化思維設置就是依據本地習俗和文化定制。當您閱讀和使用我們的重印版時，請銘記此原則，盡可能向我們提供有關您希望未來版本是何樣才能更好地幫助您的意見回應。

參考文獻：

- [1] <http://www.signalprocessingsociety.org/uploads/docs/ChineseTranslationSPMagazine.pdf>
- [2] <http://signalprocessingsociety.org/newsletter/introducing-a-translated-reprint-edition-of-ieee-signal-processing-magazine-2>

SP

Smart Grid Represents a Potentially Huge Market for the Electronics Industry

Building a global smart electric grid is an ambitious task, fraught with lots of complicated technical and public policy issues. But it's already turning out to be a big pay day for several industry companies that have committed to helping make it happen.

IEEE Signal Processing Society President Mos Kaveh has referred to the smart grid (*IEEE Signal Processing Magazine*, vol. 27, no. 3, p. 4, May 2010) as a "major initiative" with "significant signal processing content," requiring sensing, analysis, communication, and actuation technologies.

"We're seeing signs in the utility and power generation industry that suggest regulator and customer interactions and technology and business alliances will rapidly evolve as these massive infrastructure projects are executed," says David Etheridge, utilities and power generation industry leader at Price-waterhouse Coopers, LLP.

In his report, "Smart Grid Growing Pains," Etheridge says, "Utilities are forming alliances with the auto and tech industries, adding several layers of operational complexity. These factors could, in some ways, redefine the way utilities will do business going forward. The biggest question," he says, "is whether customers will embrace or reject smart grid technology."

WHAT IS THE SMART GRID?

The best and most straightforward description can probably be found in the Energy Independence and Security Act of 2007, which defines the term "smart grid" as the modernization of the electricity delivery system that monitors,

protects, and automatically optimizes the operation of its interconnected elements—from a central and distributed generator through a high-voltage transmission network and distribution system, to industrial users and building automation systems, and to end-use consumers and their thermostats, appliances, other handheld devices, and even electric vehicles.

The U.S. Government has awarded US\$4.5 billion in grants to help build a smart electric grid across the country. Smart grid pilot and demonstration programs are already under way in 33 states. China expects to allocate US\$596 billion to deploy smart grid applications. The European Parliament has approved a European Union (EU) energy package with a mandate for the roll-out of smart utility meters in all EU member states by 2022, with 80% coverage expected by 2020.

The U.S. Department of Energy (DoE) has already financed 54 projects with US\$100 million in American Recovery and Reinvestment Act of 2009 (ARRA) smart grid workforce training programs. The winners of these funds are in several states and represent collaborative efforts of utilities, universities, technical schools, and manufacturers. Part of this program is aimed at finding ways to integrate digital devices into the mechanical electric grid.

BIG ELECTRONICS BUCKS

Currently, there are about 2.7 billion utility meters in use globally, about half of which are electric. Most meters (gas, water, electric) can be read manually. Only about 8% of the meters used globally are "automatic," meaning they can be read by a utility company remotely. That's about to change as the trend now is toward automatic meter reading

(AMR), which not only increases the electronic content of the meter but drives a networked infrastructure for remote metering.

Smart meter manufacturers and the semiconductor vendors that supply them anticipate a huge market for their products as a result of the development of the smart grid. At least 500 million electric meters are expected to be replaced in the next ten years, representing a market of US\$750 million over that period of time for chip manufacturers, according to data developed from a variety of sources.

The market for communicating electricity meters in North America has already seen record growth over the past year, according to IMS Research, with 2009 shipments exceeding 8 million units, of which a third were RF mesh meters. However, 2010 is expected to see well over 7 million unit shipments of RF mesh meters alone, raising the product share of RF mesh in communicating meters to well over 50% in North America. But Europe and Asia, according to IMS Research, continue to favor high-speed power line communications meters.

Demand for electricity is growing faster than the supply infrastructure, according to "The Technology of the Smart Grid," a 140-page report published by Research and Markets. At the same time, the conversion to a digital economy is increasing the demand for high-quality power. The need to reduce peak demand and improve power quality requires an enhanced distribution grid that can both heal itself and provide two-way communication between the utility and the end-user, the report points out.

Another study by the market research organization Frost & Sullivan agrees with the need for a complete modern power grid but focuses on the European

智慧電網為電子業帶來巨大的潛在市場

建

立全球智慧電網是一項艱鉅的任務，其間充斥著大量複雜的技術和

公共政策問題。然而有數家致力於協助實現智慧電網的工業公司，已因此賺進大把鈔票。

IEEE 訊號處理學會會長 Mos Kaveh 稱智慧電網（*IEEE 訊號處理雜誌*，第 27 卷第 3 期，第 4 頁，2010 年 5 月）是具有「實義訊號處理內容」的「重大創始」，其中需要運用感測、分析、通訊和驅動技術。

普華永道會計師事務所的公用事業暨發電行業主管 David Etheridge 聲稱：「我們從公用事業和發電行業看到蛛絲馬跡，其中顯示監管機構和客戶互動以及技術和商業聯盟將迅速演變成這些已在執行的龐大基礎設施專案」。

Etheridge 在他的《智慧電網成長煩惱》報告中指出：「公用事業正在與汽車及高科技產業形成聯盟，這使得操作複雜性增加了好幾倍。在某個程度上，這些因素可能會重新定義公用事業日後的生意之道。然而最大的問題是，用戶是會接受還是拒絕智慧電網技術。」

什麼是智慧電網？

我們或許可以在 2007 年的新能源法（Energy Independence and Security Act of 2007）中找到最好且最直接的描述，此法案將「智慧電網」這個術語定義為電力輸送系統的現代化，其中涵蓋監控、保護及自動最佳化互聯元素的運作——從中央和分散式發電機開始，透過高壓電傳輸網路和配送系統，到達工業使用者和建築物自動化系統，以及終端消費者與他們的自動空調、家用電器、其他手持設備乃至於電動車輛。

美國政府已撥款 45 億美元，用以協助建立全國性智慧電網，其測試暨示範計劃已在全美 33 州展開。中國大陸則投入 5960 億美元，開始部署智慧電網設備。歐洲議會已批准歐盟能源配套計畫，其中強制所有歐盟成員國必須在 2022 年之前安裝智慧電錶，預計 2020 年將達到 80% 的涵蓋率。

美國能源部（DOE）已透過 2009 年美國復甦及再投資法案（ARRA）的智慧電網勞動力培訓計畫，為 54 個專案注資 1 億美元。這些資金獲取者分散美國各州，而且代表公用事業、大學、技術學校與製造商之間的協同合作。這項計畫的目標之一，就是尋找方法以便在機械電網中整合數位設備。

大型電子變換器

目前全球使用中的電錶數目約為 27 億，其中一半左右為電子型。大多數的（天然氣、供水、供電）電錶都是手動讀取，只有大約 8% 的電錶是「自動」的，也就是說公用事業公司可從遠端讀取。然而情況即將改變，現在的趨勢正邁向自動計量系統（AMR），這樣不僅可以增加電錶的電子內容，還可以推動基礎設施網路化以利發展遠端計量。

因為智慧電網的發展，智慧電錶製造商和半導體供應商的產品於是有了龐大市場。預計在未來十年，至少需要汰換 5 億台電錶，綜合多種資料來源的看法，這段期間將為晶片製造商創造 7.5 億美元的市場。

根據 IMS Research，北美通訊電錶市場在過去一年出現創紀錄成長，2009 年出貨量超過 800 萬台，其中三分之一是射頻（RF）網路電錶。然而，預計北美 2010 年就可以看到射頻（RF）網路電錶出貨量超過 700 萬台，如此一來在北美的通訊電錶中，射頻（RF）網路產品的市佔率就會超過 50%。但在歐洲和亞洲，根據 IMS Research 的資料，則依然看好高速電力線通訊電錶的發展。

smart grid market for advanced components. “The rise in offshore wind and solar farms has resulted in the logical move of opting for high voltage direct current for electrical power transmission, which is set to become the norm,” says Vikas Ravindran, an F&S senior research analyst.

DIGITAL SIGNAL PROCESSOR APPLICATIONS

Digital signal processors (DSPs) are already finding several applications in the development of smart grid systems.

Analog Device’s Blackfin ADSP-BF50x processors have been optimized for renewable energy and smart grid infrastructures, such as enabling inverters to convert variable dc output into “clean” current and regulate power flow into the commercial electric grid and/or local electrical networks fed by residential and wind turbines. DSPs are also used in the design of substation monitoring equipment. They’re also an option in isolation/interface metering architectures.

Analog Devices thinks about the smart grid market in terms of three segments: power generation, transmission distribution, and smart meters.

Power generation means mostly wind and solar power, which Ronn Kliger, director of Analog Devices’ Energy Group, says opens up a significant demand for DSP technology. Wind is actually the big winner here because the signal processing technology required is more sophisticated and complex than for solar. Optimizing the angle of the blades of wind mills is just one example of the challenge.

The processing requirements are quite as complex for solar systems and DSPs will find a place in solar. But there’s also plenty of room for microcontrollers (MCUs), particularly for lower-end systems. “The trend is not one of going to higher processing power levels, but moving towards integrating some analog functions into the processor, regardless of whether it’s an MCU or DSP.”

In the transmission and distribution area, most of the chips are going into substations. The primary application here for the smart grid is to monitor the

quality of the voltages and currents on the line and to communicate that information to a central control location, but also be able to take action based on such issues as the detection of a fault condition. This sector is becoming a major user of DSPs because of the amount and sophistication of processing required. For example, substations can use dozens of lines that need to be monitored, and each one may require harmonic analysis, calculating voltages, and currents. Transmission and distribution also require communications, but in this area, DSP usually gives way to high-end MCUs.

“We’re starting to get requests to integrate these functions,” says Kliger. “That is, to come out with processors with both the signal processing capability to support the analysis but also to support the communications capability.”

ANALOG DEVICES THINKS ABOUT THE SMART GRID MARKET IN TERMS OF THREE SEGMENTS: POWER GENERATION, TRANSMISSION DISTRIBUTION, AND SMART METERS.

The smart meter case is more interesting for the digital signal processing community. For one thing, meters don’t usually need the processing power offered by a DSP. Where DSPs tend to fit into this sector, at least for Analog Devices, is in its metrology products, which Kliger says isn’t always transparent to meter designers and manufacturers. With a metrology approach, an A/D converter is used with the outputs from that feed to an MCU, which the meter manufacturer would have to program. In fact, because communications standards vary globally, chip suppliers have developed custom, programmable chips to meet the requirements of certain meter manufacturers. The penalty for this approach is cost. Some meter makers are dealing with this by focusing their marketing strategy on only a few countries—at least until the standards issue is sorted out. But this could easily

be complicated by the EU’s requirement that most European consumers adopt smart meters by 2020.

TEXAS INSTRUMENTS LOOKS TO MCUs

Texas Instruments (TI) is another major player in the smart grid arena but believes that low power wireless communication for electricity meters opens the market to non-DSP solutions.

To meet legislative and standards requirements both in the United States and Europe, TI has added at least 16 new ultra-low power MCUs to its electricity and gas metering portfolio. TI is also targeting power line communications (PLCs) and radio frequency (RF) interfaces for advanced metering infrastructures expected to be a significant part of smart grids around the world.

In May, TI introduced its first 2.5-A step-down SWIFT converter to support input voltages up to 60 V for Global System for Mobile Communications and General Packet Radio Service (GPRS) modules used in electronic meters, fleet management, and security systems. The new TPS54260 monolithic synchronous switcher with integrated high-side field-effect transistor (FET) provides up to a 95% efficiency and low operating current to 12-, 24-, and 48-V point of load designs.

TI also recently bolstered its analog manufacturing capacity with the purchase of more than 100 tools from Qimonda North America and Qimonda Dresden, Germany. The company says this is a first step in launching the phase II expansion of RFAB, the industry’s first 300-mm analog wafer fab, located in Richardson, Texas.

Olivier Monnier, TI’s metering marketing manager, says the smart grid falls into the company’s overall energy initiative, where TI is conducting several R&D programs based on power, solar, and lighting. At least eight programs are linked specifically to the smart grid. “Much of the focus is on enhancing communications techniques for the smart grid, including next-generation electrical meters [smart meters],” notes

根據 Research and Markets 出版的長達 140 頁的《智慧電網技術》報告，電力需求的增長速度比供電基礎設施的發展還快。在此同時，轉向數位經濟的結果將會增加高品質電力需求。該報告指出，減少高峰用電需求與改善供電品質不但需要強化輸電網的自我復原能力，還需要提供公用事業與最終用戶之間的雙向溝通。

市場研究機構 Frost & Sullivan 所提出的另一份研究也同意確實存在完全現代電網需求，只不過其重點放在歐洲智慧電網市場的先進元件。F & S 的資深研究分析師 Vikas Ravindran 表示：「海上風力發電和太陽能發電場的興起，必然會轉向選擇以高壓直流電進行電力傳輸，而這將會成為常態。」

數位訊號處理器應用

數位訊號處理器（DSP）在智慧電網系統的發展中已有諸多應用。

Analog Device 公司的 Blackfin ADSP - BF50x 已針對處理器可再生能源和智慧電網基礎設施進行最佳化，如通過逆變器將可變直流輸出轉換成「乾淨」的電流，並調節進入商業電網的能流和/或由住宅和風力渦輪機供電的局部電網。在變電站監控設備的設計中，也會使用數位訊號處理器。此外，在隔離/連接計量架構中，數位訊號處理器也是一個選項。

Analog Device 認為可以從三方面來考量智慧電網市場：發電、輸配電和智慧電錶。

發電方式主要是風力和太陽能發電，ADI 公司能源部門總監 Ronn Klinger 稱這種情況造就了龐大的 DSP 技術需求。其實風力是最大贏家，因為其所使用的訊號處理技術比太陽能更精細更複雜，風車葉片角度最佳化就是一個頗具挑戰的例子。

太陽能系統的處理需求也相當複雜，DSP 將會在其中謀得一席之地。不過，微控制器（MCU）的發展空間也很大，特別是在低階系統。「無論是微控制器還是 DSP，趨勢走向不是朝更高的處理能力推進，而是要在處理器中整合某些類比功能。」

在輸配電領域，大部分晶片都位於變電站。智慧電網在此的主要應用是監視輸電線上的電壓和電流的品質，然後將該資訊傳送給中央控制位置，同時還可以在偵測到故障時根據問題狀況採取行動。這部分因為需要大量且精細的處理，所以成為 DSP 的主要用戶。例如，變電站會用到數十條都需要檢測的輸電線，其中每一條都需要諧波分析、計算電壓和電流。

ANALOG DEVICES 認為可以從三方面來思考智慧電網市場：發電、輸配電和智慧電錶。

輸電和配電也需要通訊，但在這方面數位訊號處理器通常不敵高階微控制器。

Klinger 表示：「我們開始要求整合這些功能。也就是說，將推出具有兩種訊號處理能力的處理器，除了能夠支援分析之外，還要能夠支援通訊能力。」

對數位訊號處理社群來說，智慧電錶的狀況就更為有趣。首先，電錶通常不需要單個 DSP 的處理能力。多個 DSP 才能應付這部分的狀況，至少在 Analog Devices 是如此，Klinger 提到的計量產品部分對於電錶設計工程師和製造商來說並不總是那麼透明。在計量方法中，A/D 轉換器將來自電錶的輸出饋送給 MCU，而電錶製造商不得對其進行程式設計。事實上，由於全球各地的通訊標準不盡相同，晶片供應商已開發出可程式化的自訂

晶片，以滿足特定電錶製造商的需求。採取這種做法的不利之處就是成本。為了解決此問題，有些電錶製造商只集中在少數國家進行銷售，在解決通訊標準問題之前這種狀況會保持下去。不過，歐盟要求大部分的歐洲消費者須在 2020 年前採用智慧電錶，這使得狀況更為複雜。

德州儀器的 MCU 展望

德州儀器（TI）是智慧電網領域的另一個要角，但是他們相信，低功率無線通訊才是電錶市場的王道，而非 DSP 解決方案。

為滿足美國和歐洲的法律和標準要求，TI 在他們的電力 and 燃氣計量產品系列中新增了至少 16 種超低功率 MCU。此外，TI 還聚焦於先進計量基礎設施的電力線通訊（PLC）和射頻（RF）介面，這兩項預計將成為全球智慧電網的重要組成部分。

今年五月，TI 推出首款 2.5A 步降 SWIFT 轉換器，支援高達 60V 的輸入電壓，適用於智慧電錶、車隊管理和安全系統中的全球行動通訊系統和通用分組無線業務（GPRS）模組。新的 TPS54260 單片同步開關轉換器集成有高壓側場效應電晶體（FET），效率高達 95%，且可為 12V、24V 和 48V 的點荷載設計提供較低的工作電流。

TI 近期還向 Qimonda North America 公司和德國 Qimonda Dresden 公司購買 100 多種工具，藉此強化自身的類比製造能力。該公司表示，這是啟動 RFAB 第二階段擴張的第一步，其首個 300 mm 類比晶圓廠設於美國德州的 Richardson。

TI 的計量行銷經理 Olivier Monnier 聲稱，智慧電網已納入該公司的整體能源新方案，TI 正在展開多項發電、太陽能和照明研發專案，其中至少有八個專案明確地與智慧電網相

Monnier. Home displays are another area of interest of TI, he says.

Looking at these applications, TI sees opportunities in MCUs, or application-specific embedded processing (low-power MCUs for metrology, Cortex M3 MCUs for application processors, MCUs for the PLC, and low-power RF SoCs), for low-power RF, power management, and radio-frequency identification (RFID). Monnier says TI sees additional opportunities in substation applications, where data sent from meters are concentrated—typically a DSP-type or processor engine function.

To measure energy from smart meters, TI is going with its low-power MCU MSP430, which is already used with utility meters and in-home applications. But the new meters must also handle some new requirements for utilities. One is “time of usage,” which implies the need for more memory, connectivity to multiple communications options, and security. TI’s response at the moment is its Stellaris Cortex M3 platform, built around a main application processor for handling typical MCU tasks.

STANDARDS AND WIRELESS

Wireless is expected to be the dominant communications technology for the smart grid in the United States. In fact, wireless carriers and system vendors are already carving out a role for themselves in this market.

IEEE Standards Association (IEEE-SA) officially formed its own smart grid initiative in May 2009, focusing on power engineering, communications and IT, and launched a project to create *The IEEE Standard 2030 Guide for Smart Grid Interoperability of Energy Technology and Information Technology Operation with the Electric Power System (EPS) and End-Use Applications and Loads*.

The guide (IEEE-SA P2030) is expected to provide a knowledge framework for understanding the defining smart grid interoperability of the electric power system with end-use applications. It’s designed to set the stage for future smart grid standards.

IEEE-SA is one of several groups collaborating with the National Institute of Standards and Technology (NIST) to create a “Smart Grid Interoperability Standards Roadmap,” a document that will identify the short- and long-term plans for architecture development and infrastructure development for the smart grid.

IEEE has more than 100 standards published and in development that are potentially crucial to the smart grid that span digital information and controls technology, networking, security, reliability, assessment, interconnection of distributed resources, sensors, electric metering, broadband-over-power-line, and systems engineering.

One possible new standard, IEEE 802.15.4G, that covers smart utility networks, recently went out for first review and is expected to have an impact on

THE IEEE HAS ALREADY HELD SEVERAL CONFERENCES ON SMART GRID ISSUES, SUCH AS THE IEEE GREEN TECHNOLOGY CONFERENCE, IN APRIL IN GRAPEVINE.

wireless smart meter service. (One of the issues covered in this standard proposal is modulation technique, which directly impacts processing choices, including DSPs.)

The IEEE has already held several conferences on smart grid issues, such as the IEEE Green Technology Conference, in April in Grapevine, Texas. Additional conferences, Webinars, and public outreach forums in smart grid technology and public policy are planned. The IEEE has also published nearly 2,500 papers on smart grid issues in more than 40 journals and is launching two new cross-disciplinary and archival journals, *IEEE Transactions on Smart Grid* and *IEEE Transactions on Sustainable Energy*.

ZIGBEE AND NIST

The ZigBee Alliance is another key player that is driving the development of smart grid wireless technology. With

more than 350 industry members, the organization made its ZigBee’s Smart Energy Profile 2.0 available for public download and comment in late April. The software package enables service providers and utilities to wirelessly communicate with and manage common household devices such as smart thermostats and in-home displays and appliances.

“This is an integral part of home area networks (HAN) and the emerging requirements for smart grid interoperability and energy management,” says George Arnold, national coordinator for Smart Grid Interoperability at NIST, which included the ZigBee Smart Energy 2.0 in its *NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 1.0* in January. Common HAN devices include utility meters, thermostats, pool pumps, water heaters, appliances, and plug-in electric vehicles.

The ZigBee Alliance has also agreed to collaborate with IP for Smart Objects (IPSO) on wireless HANs using ZigBee’s IP specification and Smart Energy Version 2.0 standard.

RF Micro Devices, Inc., a major producer of chips in several semiconductor process technologies (but not DSPs) has teamed with Ember Corp., which develops wireless mesh networking technologies, to introduce ZigBee front-end modules that will give utilities and consumers more control over how they monitor and save energy.

General Electric (GE) has launched its own pilot program in the United States to run smart grid applications over a worldwide interoperability for microwave access (WiMAX) network. GE expects the trial, conducted by Michigan-based utility Consumer Energy, to demonstrate a number of benefits, including increased efficiency, more bandwidth, security, lower costs, and the ability to react to potential outages before they occur. GE also has a US\$3.3 million grant from the U.S. DoE to work with Arizona Public Service to determine the best way to integrate large amounts of solar power into the existing grid.

關。Monnier 強調：「注意力大多放在加強智慧電網的通訊技術，包括下一代電錶 [智慧電錶]。」他還表示，室內顯示也是 TI 感興趣的另一個技術領域。

目睹這些應用，TI 在用於低功耗 RF，電源管理和射頻識別 (RFID) 的 MCU 或特定應用的嵌入式處理 (計量用低功耗 MCU，應用處理器的 Cortex M3 MCU，PLC 用 MCU，以及低功耗射頻系統單晶片) 中看到了機會。Monnier 聲稱，TI 在變電站的應用中看到了更多的機會，其中電錶發送的資料較密集，通常使用 DSP 型或處理器引擎功能。

要測量來自智慧電錶的能量，TI 使用其低功耗 MCU MSP430，此產品早已用於需給電錶和家庭應用。不過，此新電錶還必須處理一些新出現的需給要求。其中之一就是“用時”，這意味著需要更多的記憶體、多種通訊系統連線選項與安全性。TI 對此問題的回應是目前他們的 Stellaris Cortex M3 平台，此平台建立在用於處理典型 MCU 作業的主要應用處理器之上。

標準和無線

無線通訊預計將會成為美國智慧電網的主要通訊技術。事實上，無線運營商和系統供應商已在這個市場為自己找到一席之地。

IEEE 標準協會 (IEEE - SA) 在 2009 年 5 月正式擬定自己的智慧電網方案，其中以電力工程、通訊和 IT 為重點，另外還啟動一個專案以建立 IEEE 標準 2030 指南：能源技術與資訊科技、電力系統 (EPS) 以及終端應用與負載的智慧電網互通性。

此指南 (IEEE- SA P2030) 旨在提供知識架構，以理解電力系統與終端應用的智慧電網互通性定義；其設計目的是，為未來的智慧電網標準奠定基礎。

IEEE - SA 是和美國國家標準與技術研究院 (NIST) 合作建立「智慧電網互通性標準藍圖」的數個團體之一，此文件確立智慧電網架構開發與基礎設施發展的短期及長期計畫。

IEEE 已經發佈或正在開發的標準多達 100 多個，這些標準對於橫跨數位資訊和控制技術、網路、安全、可靠性、評量、分散式資源互連、感測器、電能計量、電力線寬頻通信與系統工程的智慧電網而言是潛在關鍵。

涵蓋智慧公用事業網路且可能成為新標準的 IEEE802.15.4G，最近已發佈

IEEE 已就智慧電網問題舉行數次會議，例如四月份在 GRAPEVINE 舉行的 IEEE 綠色技術會議。

並接受第一次審查，預計將會對無線智慧電錶服務造成影響。(本標準提案所涉及的問題之一是調製技術，它會直接影響處理過程的選擇，包括 DSP。)

IEEE 已就智慧電網問題舉行了多次會議，如 4 月份在美國德州 Grapevine 舉行的 IEEE 綠色技術會議。目前正在醞釀有關智慧電網技術與公共政策的其他會議、網路研討會與公共宣傳論壇。IEEE 還在 40 多個期刊中出版近 2500 篇有關智慧電網問題的論文，另外還推出兩本全新的跨學科存檔期刊：IEEE 智慧電網彙刊 (IEEE Transactions on Smart Grid) 和 IEEE 永續能源彙刊 (IEEE Transactions on Sustainable Energy)。

ZIGBEE 和 NIST

在推動智慧電網無線技術發展方面，ZigBee 聯盟是另一個要角。該聯盟共有 350 多個業界成員，他們在四月底製作了供大眾下載和發表評論的「ZigBee 智慧能源概述 2.0」。該套裝軟體能夠讓服務提供者和公用事業公司與常見的家用電器進行無線通訊，並對其進行管理，例如智慧恆溫器、家用顯示器和電器設備。

NIST 的智慧電網互通性全國協調者 George Arnold 表示：「這是家庭區域網路 (HAN) 不可或缺之部分，也是智慧電網互通性與能源管理的新興需求。」在一月份發佈的 NIST 智慧電網互通性標準的架構與藍圖 1.0 版中包含 ZigBee 智慧能源 2.0。常見的家庭區域網路設備包括公用事業電錶、空調、抽水機、熱水器、家用電器和插電式電動車。

ZigBee 聯盟還同意與 IPSO (IP for Smart Objects) 合作，在無線家庭區域網路中使用 ZigBee 的 IP 規格與標準的智慧能源 2.0 版標準。

作為擁有多個主要半導體工藝技術 (並非 DSP) 的晶片製造商，RF Micro Devices 公司與 Ember 公司聯手開發無線網狀網路技術，其所推出的 ZigBee 前端模組可讓公用事業公司與消費者增進監測控制並且節省能源。

通用電氣公司 (GE) 已經在美國推出自己的試點計劃，他們要透過全球互通微波存取 (WiMAX) 網路進行智慧電網應用。GE 希望透過此測試 (由總部位於密西根的公用事業 Consumer Energy 執行)，向人們展示諸多好處，包括提高效率、更寬的頻寬、安全性和較低的成本，還有能夠在潛在故障發生之前做出應對。GE 還從美國能源部獲得 330 萬美元，與

Networking specialist Cisco Systems, which has demonstrated a growing interest in consumer electronics, has made an equity investment in GridNet, a smart metering company, while Sprint, the wireless carrier, which has partnered with both GE and GridNet, is promoting WiMAX for use in the smart grid. Another WiMAX vendor, Alvarion, has been conducting pilot studies with several utilities.

Honeywell believes it has boosted its opportunities in the smart grid market with the recent acquisition of San Rafael, California-based Akuacom, a leader in automated demand response technology and services for the smart grid. Akuacom's demand response automation server provides utilities and independent system operators with two-way communication with energy management systems at commercial and industrial sites. Honeywell says the system fits well into its already developed energy management and smart grid portfolio.

THE POWER LINE PLAY

One of the biggest issues is developing a power line standard for the smart grid. NIST has been reviewing industry input as part of its task of writing a first draft for an overall smart grid standard and is working closely with the Smart Grid Interoperability Panel (SGIP), which has 15 priority action plan committees working on several smart grid-related standards, including power lines.

Stakeholders involved in this effort have produced some lively debate over standards options, but one possibility that is being seriously considered for adoption is IEEE 1901, a standard that covers multiple power line approaches.

The HomeGrid Forum has been formed as a global trade group specifically to develop next-generation, unified coaxial phone line and global power line home networking standards and to promote their adoption. HomeGrid has been working with the International Telecommunications Union's (ITU-T) G.hn compliance and interoperability working group. The

ITU recently commissioned G.hnem, a new project within G.hn, to address home networking electric vehicle communications and grid access aspects of energy management, based largely on existing low-frequency narrowband orthogonal frequency-division multiplexing technologies.

"G.hn and G.hnem have the ability to connect hundreds of end products in and outside of the home, allowing manufacturers and service providers to overcome traditional challenges while enabling new functionality," says

ONE OF THE BIGGEST ISSUES IS DEVELOPING A POWER LINE STANDARD FOR THE SMART GRID.

Matthew Theall, president of the HomeGrid Forum and a technology strategist at Intel executive.

In June, the G.hn standard for next-generation wired home networks was approved by the United Nations' ITU-T. All three components of the new standard—the Data Link Layer (G.9961), Physical Layer (G.9960), and the coexistence protocol (G.9972)—are now complete. Several manufacturers are expected to bring semiconductors for multimedia home networking to market later this year. (Kawasaki Microelectronics America, Inc. and TangoTec are the newest members of HomeGrid Forum, joining chip suppliers DS2, Ikanos, Intel, Lantiq, Sigma Designs, and TI as backers of the newly approved standard.)

"G.hn is the only technology platform that supports every wire in the home," says Theall. "We believe that consumer electronics manufacturers, PC companies, and service providers and every group that works within these markets will benefit from the commercial availability of this technology as it rolls out later this year."

The Association of Home Appliance Manufacturers (AHAM) has enlisted EnerNex Corp., a consulting firm that is also helping NIST organize smart grid standards efforts, to draft a detailed review of all home network technologies.

AHAM says it expects the review to help it decide how to advise members on smart grid issues.

THE EUROPEAN MARKET

Globally, British Gas has placed a multi-million pound contract with Vodafone to enable smart metering services that will use GPRS to backhaul usage data. The contract will see Vodafone supply nearly a million GPRS connections to household meters to help monitor real-time energy usage. Consumers will be able to monitor their electricity and gas usage via an in-home display. Landis+Gyr, a provider of integrated energy management solutions, will develop the smart meters and a touch-screen hub (see Figure 1.) Trilliant, a smart meter networking specialist, which purchased the WiFi mesh network vendor SkyPilot in May 2009, will connect the various units. OSIsoft has been signed to store data, and SAP will integrate the system with British Gas's existing IT systems. Landis+Gyr is also deploying an advanced metering system to more than 700,000 AEP Texas electric customers in Texas.

Siemens, another major player in the European market, uses Analog Devices' Blackfin DSPs in its metering products to handle continuously variable wide-band frequency signals.

"Our smart metering solution is the result of many years of research," says Alexander Schenk, business segment manager for Automated Metering and Information Systems (AMIS) at Siemens



[FIG1] Landis+Gyr is developing smart meters similar to the one shown here for several markets, including Europe. (Image courtesy of Landis+Gyr.)

Arizona 公共服務部門共同合作，以判斷將大量太陽能發電融入現有電網的最佳方式。

網路專家 Cisco Systems 公司對消費電子產品的興趣越來越濃厚，他們已對智慧電錶公司 GridNet 進行股權投資，而已與 GE 和 GridNet 合作的無線電信廠商 Sprint，正在推動 WiMAX 在智慧電網中的應用。另一家 WiMAX 廠商，Alvarion，也與多家公用事業公司合作展開測試研究。

Honeywell 相信自己在智慧電網市場發展機會大增，因為他們近期收購總部位於加州 San Rafael 的 Akuacom 公司，該公司是智慧電網自動化需求回應技術與服務方面的佼佼者。Akuacom 的需求回應自動化伺服器可在商業和工業場合，為公用事業公司和獨立系統營運商提供與能源管理系統之間的雙向通訊。Honeywell 表示，該系統非常適合他們已經開發的能源管理和智慧電網系列產品。

電力線的影響

最大的問題之一是開發智慧電網的電力線標準。NIST 在撰寫整個智慧電網標準初稿時，其中部分工作就是檢討行業投入，另外還與智慧電網互通性委員會 (SGIP) 緊密合作，其中有 15 個優先行動計畫委員會正在緊鑼密鼓地制定數個與智慧電網密切相關的標準，其中包括電力線。

參與這項工作的利害相關人士就標準選擇進行激烈辯論，其中大家認真考慮採用的是 IEEE1901 標準，其中涵蓋多種電力線方式。

HomeGrid 論壇已經成為全球性質貿易團體，專門致力開發新一代統一的同軸電話線與全世界通用的電力線家庭網路標準，以及促進這些標準的應用。HomeGrid 一直在推動國際電信聯盟 (ITU - T) 的 G.hn 規範與互通

性工作小組。ITU 最近啓用 G.hnem，這是隸屬於 G.hn 的一個新專案，目標是解決家庭網路電動車輛通訊及能源管理的電網存取問題，其中主要採用低頻窄頻或正交頻分複用技術。

HomeGrid 論壇主席暨 Intel 高層技術戰略專家 Matthew Theall 表示：「G.hn 和 G.hnem 能夠連接家庭內外數以百計的終端產品，讓製造商和服

最大的問題之一是開發智慧電網的電力線標準。

務供應商能夠克服傳統挑戰，同時啓用新功能。」

聯合國 ITU-T 在六月份批准了適用於下一代家庭有線網路標準 G.hn。新標準的三個組成部分：資料連結層 (G.9961)、實體層 (G.9960) 和共存協定 (G.9972) 目前均已完成。在今年稍後，預計市場上將會有各家製造商推出多媒體家庭網路適用的半導體。(Kawasaki Microelectronics America 公司和 TangoTec 是 HomeGrid 論壇的新成員，他們與晶片供應商 DS2、Ikanos、Intel、Lantiq、Sigma Designs 和 TI 同為新近核准標準的支持者。)

Theall 表示：「G.hn 是唯一支援所有家庭纜線的技術平台。我們認為，當這項技術在今年稍後上市時，消費電子產品製造商、PC 廠商、服務提供商，以及在這些市場中營運的所有團體，都可從本技術的商業化受惠。」

家電製造商協會 (AHAM) 獲得 EnerNex 顧問公司的協助 (該公司也協助 NIST 管理智慧電網標準)，以針對所有家庭網路技術起草詳細的評論報告。AHAM 表示，希望他們的評估報告有助於決定，該如何就智慧電網問題向成員提出建議。

歐洲市場

從全球來看，英國天然氣公司與 Vodafone 簽訂金額高達數百萬英鎊的智慧計量服務啓用合約，其中使用 GPRS 回傳使用資料。該合約要求 Vodafone 提供近百萬的家用電錶 GPRS 連接，以監測即時能源使用情況。消費者可以透過室內顯示器，監測其電力及天然氣的使用狀況。綜合能源管理解決方案供應商 Landis+Gyr，即將開發智慧電錶和觸控式螢幕集線器 (見圖 1)。智慧電錶網路專家 Trilliant 在 2009 年 5 月收購 WiFi 網狀網路供應商 SkyPilot，Trilliant 將會連接各種不同設備。SAP 已經與 OSIsoft 公司簽約以利儲存資料，SAP 將會整合該系統與英國天然氣公司現有的 IT 系統。Landis+Gyr 則正在美國德州為 70 多萬 AEP Texas 電力用戶部署先進的計量系統。

歐洲市場的另一個要角 Siemens，他們在計量產品中使用 ADI 的 Blackfin DSP，以處理寬頻連續可變頻率訊號。

「西門子能源配電部門自動化計量和資訊系統 (AMIS) 的業務分部經理 Alexander Schenk 表示：「我們的智慧電錶解決方案是多年研究的成果。」(西門子行銷長 Tom



【圖1】Landis+Gyr 正在為多個市場開發與此類似的智慧電錶，其中包括歐洲。(圖片來自 Landis+Gyr)

WHAT'S A SMART GRID?

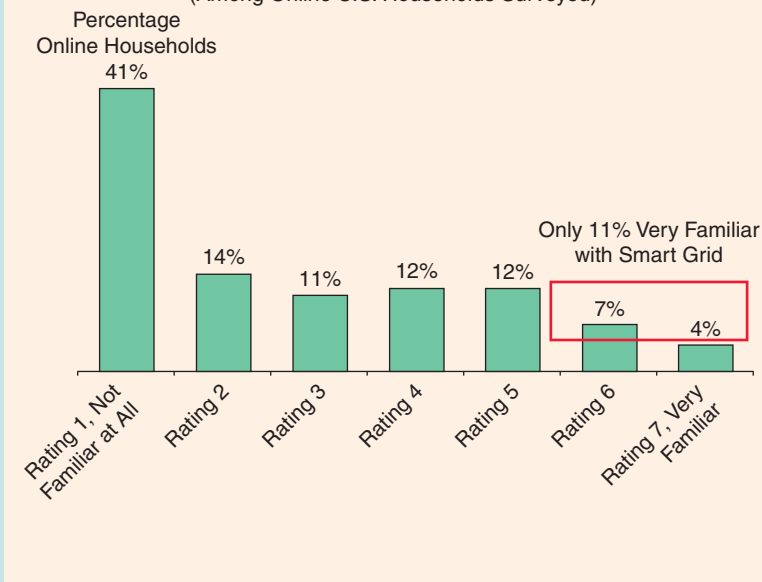
Utilities have already installed 13 million smart meters in U.S. households, but a recent survey by Parks Associates, a market research company specializing in tracking trends in consumer technologies and services, has found consumer knowledge of smart grid technology to be seriously lacking (see Figure S1). In fact, according to a Parks report, "Residential Energy Management Opportunities for Digital Systems and Services," only 11% of U.S. consumers are familiar with the term "smart grid."

Parks believes this could pose a risk to future deployment plans of utilities by opening the door for communications system providers to offer their own energy monitoring solutions. (Potentially bad news for utilities since the Parks study also found that just over 80% of households said they are willing to pay US\$80–\$100 for cost-saving equipment if they are guaranteed to save 10–30% off their monthly electricity bill.) Parks also notes that more than 80% of consumers surveyed say they are very interested in learning how to cut their energy costs—even though less than half indicated a desire to learn more about the smart grid.

"Utilities plan to deploy more than 50 million smart meters by 2014, but they need to improve their educational outreach if they want consumers to embrace this technology," says Bill Ablondi, director of home systems research at Parks Associates. "Growing skepticism of the efficacy of smart meters could also hinder growth." Still, Ablondi says that close to 50% of U.S. households are interested in home energy monitoring, creating new opportunities for home controls vendors and their suppliers.

In fact, utilities are already feeling some heat from consumers from the backlash of lawsuits filed in Texas and

How Familiar Are You with the Term "Smart Grid?"
(Among Online U.S. Households Surveyed)



[FIGS1] According to the market research firm Parks Associates, only 11% of U.S. consumers are familiar with the term "smart grid," which it says doesn't bode well for utilities or others hoping to install millions of so-called smart meters in homes and business facilities. (Used with permission from Parks Associates.)

California, blaming smart meters for hikes in their electricity bills. Ablondi suggests this potentially could disrupt utility company plans to create an interconnected grid that offers advanced services such as monitoring and time-of-use billing.

"Communications service providers, which have existing customer relationships, could step in with meter-independent services," Ablondi says. "Smart grid deployment is forcing utilities to become evangelists for new technologies, which is an unfamiliar role for most of them."

Energy Power Distribution Division. (Tom Haas, Siemens' chief marketing officer, recently told a meeting of the IEEE Media staff that his company was conducting a six-city U.S. smart grid tour to promote the smart grid as one of the company's "hot spot" markets.)

STMicroelectronics is also a very active player in the smart meter market, supplying its 0.18- μ m BCD8 StarGRID DSP technology. Most recently, STM provided a set of components to Enel, Italy's largest power company for new electronic power meters destined for

distribution in Spain. The power meters are key elements of the new remote management system to be installed by Endesa, the largest electricity supplier in Spain.

For this application, STM is providing a power-line communication system-on-a-chip (SoC), its ST758x, a 32-b microcontroller (STM32), and a power-supply device, as well as metal-oxide-semiconductor FET and electrically erasable programmable read-only memory devices, as the basis of the new metering system for Endesa. Endesa

expects to complete the replacement of more than 13 million traditional electricity meters, installed in customers' houses, with its new generation meters between this year and 2015.

One of the big problems facing European vendors, sources point out, is that they're working in a very fragmented market. This is true especially in communications protocols, which differ from country to country. The technical requirement for smart metering for consumers across the EU is not completely clear at this point.

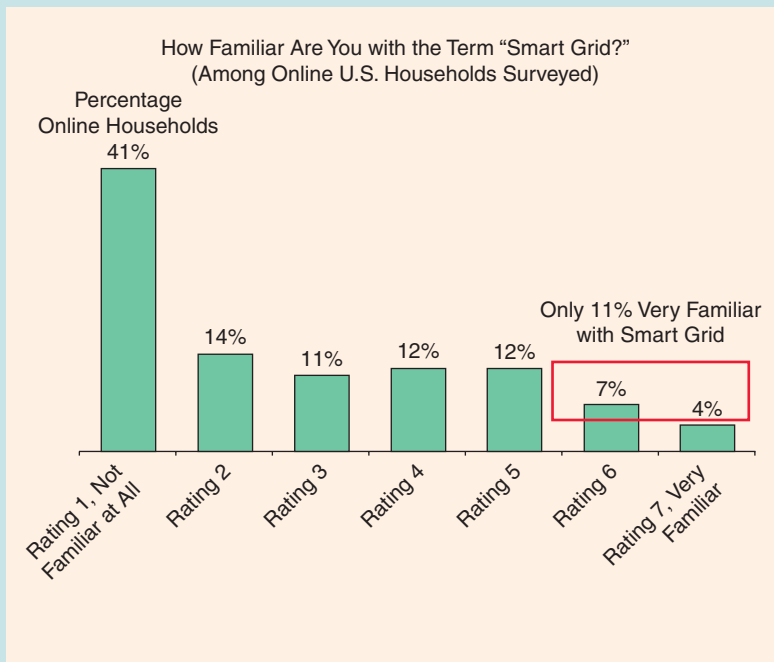
何謂智慧電網?

公用事業已經為美國家庭安裝 1300 萬個智慧電錶，但最近擅長追蹤消費者科技和服務趨勢的市場研究公司 Parks Associates 其調查研究發現，消費者嚴重欠缺智慧電網相關技術知識（見圖 S1）。事實上，根據 Parks 的報告：「住宅能源管理為數位系統和服務提供大好機會。」只有 11% 的美國消費者熟悉「智慧電網」一詞。

Parks 認為，對於有意讓通訊系統供應商提供其自家能源監控解決方案的公用事業來說，上述研究發現讓他們的未來部署計畫面臨風險。（Parks 的研究還發現，約有 80% 的家庭表示，如果能保證每月節省電費 10% 到 30%，他們願意為支付 80 至 100 美元購買節費設備，這對公用事業部門來說可能是壞消息。）Parks 還指出，超過 80% 的受訪消費者表示，他們非常有興趣學習如何削減能源費用，但其中有意進一步瞭解智慧電網的人卻不到一半。

Parks Associates 的家庭系統研究主管 Bill Ablondi 表示：「公用事業計劃在 2014 年之前部署超過 5000 萬台智慧電錶，如果他們希望消費者接受此技術，他們需要改善推廣教育。對智慧電錶效益抱持懷疑的聲浪越來越高，這也有礙智慧電錶的發展。」儘管如此，Ablondi 表示，近 50% 的美國家庭對家庭能源監測感興趣，這為家庭控制器廠商及其供應商開創了新商機。

事實上，公用事業已經感受到德州和加州消費者提起訴訟的反彈力量，這些消費者指責智慧電錶導致其電費上漲。Ablondi 認為，這可能會干擾公用事業公司的未來計劃，他們希望建立



【圖S1】根據市場研究公司 Parks Associates 的調查，僅有 11% 的美國消費者熟悉「智慧電網」一詞，該公司表示，對於希望在家庭和商業場所安裝數以百萬計所謂智慧電錶的公用事業部門或其他供應商來說，這並不是一個好兆頭。（圖片使用獲得 Parks Associates 許可。）

可提供諸如監測和使用時間計費等進階服務的互聯網。

Ablondi 表示：「現在已擁有客戶關係的通訊服務供應商，可以涉足無關電錶的服務。智慧電網的部署迫使公用事業成為新科技傳播者，然而這是他們當中大部分人最不熟悉的角色。」

Haas，他在近日會晤 IEEE Media 人員時提到，他們公司正在六城市舉行美國智慧電網巡迴活動，以促進智慧電網成為該公司的「熱門」市場之一。）

STMicroelectronics（意法半導體）也是智慧電錶市場的活躍參與者，他們提供 0.18 μm BCD8 StarGRID DSP 技術。最近，STM 提供一套元件給義大利最大的電力公司 Enel，以供要在西班牙境內配置的新型電子式電錶使用。對於西班牙最大電力供應商 Endesa 公司所要安裝的新型遠端管理

系統來說，該電錶是其中關鍵要素。

對於此應用，STM 提供電力線通訊系統單晶片（SoC）ST758x、32 位元微控制器（STM32）和電源設備，以及金屬氧化半導體 FET 和電子可改寫式可編程唯讀記憶體，以這些作為 Endesa 新型計量系統的基礎。Endesa 公司預計在今年到 2015 年之間，使用其新一代電錶替換安裝在客戶家中的 1300 多萬個傳統電錶。

消息來源指出，歐洲廠商面臨的重大問題之一是，他們在高度零散的市場中營運。在通訊協定方面尤其如

此，各國使用不同的通訊協定。對於整個歐盟的消費者而言，智慧計量的技術要求目前尚不明確。

特殊利益團體

鼓勵成員開發智慧電網的組織絡繹不絕。

IEEE 成立了智慧電網專責小組，並且積極參與 SGIP，SGIP 是一個開放透明的組織，成員涵蓋 20 多種利害關係人類別，其中包括公用事業、標準制定機構、專業社團、可再生能源產品、製造商和供應商、消費者，以

SPECIAL INTEREST GROUPS

There's no shortage of organizations promoting their members' interests in the development of the smart grid.

The IEEE has formed a Smart Grid Task Force and already has strong representation on the SGIP, an open, transparent body comprising more than 20 stakeholder categories that includes utilities, standards development organizations, professionals societies, renewable power products, manufacturers and vendors, consumers, and IT and system developers, that are expected to heavily influence the development of interoperability standards for the smart grid. Formed in November 2009, the SGIP has about 600 member organizations.

The IEEE launched the IEEE Smart Grid Web Portal (<http://smartgrid.ieee.org>) in January as an integrated gateway to smart grid information from IEEE and other expert sources. The Web site has been designed for manufacturers, policy makers, educators, academics, governments, engineers, computer scientist, researchers, and other stakeholders in the power and energy, IT, and communications industries. "Contributions from across the power and energy, communications, and information technology (IT) industries, as well as government and academia, are needed to ensure successful implementation of smart grid throughout the world," Wanda Reder, chair of the IEEE Smart Grid Task Force, said at the time of the launch of the Web site. Also, the IEEE Power & Energy Society (PES) has been extremely active in smart grid activities on several levels.

Another group, the GridWise Alliance, a technology-neutral smart grid coalition with 125 member organizations, has signed an agreement with the IEEE PES enabling them to work together on smart grid-related issues of interest to both organizations.

"The agreement to collaborate will strengthen the engineering basis of our group," says Katherine Hamilton, president of GridWise Alliance.

The GridWise Alliance has also helped launch the Smart Grid Consumer Coalition to focus on educating consumers on how they can benefit from the smart grid (see "What's a Smart Grid?").

On the consumer side, the Smart Grid Consumer Collaborative will focus on promoting the consumer adoption of the smart grid. Early members include GE, IBM, the GridWise Alliance, Magnolia/Best Buy, Ember Corp., the National Renewable Energy Laboratory, and others.

With a US\$165 billion industry at stake, the Consumer Electronics Association (CEA) has formed a Smart Grid Working Group to serve as an information exchange forum for Smart Grid projects and policies affecting the consumer electronics industry. The CEA says it will provide input to both the SGIP and NIST.

AHAM has enlisted EnerNex Corp., a consulting firm that is also helping NIST organize smart grid standards efforts, to draft a detailed review of all home network technologies. AHAM says it expects the review to help it decide how to advise members on smart grid issues.

The IEEE has scheduled IEEE SmartGridComm 2010 at NIST's offices in Gaithersburg, Maryland, 4-6 October 2010. The conference will focus on the communications aspects of the smart grid with researchers from academia, industry, and the national labs expected make presentations and exchange ideas during the conference. Additional information on the conference is available at <http://www.ieee-smartgridcomm.org>.

SP

及 IT 和系統開發商，這些成員對於智慧電網互通性標準的發展極具影響力。SGIP 成立於 2009 年 11 月，擁有大約 600 個成員組織。

IEEE 在一月份推出 IEEE 智慧電網入口網站 (<http://smartgrid.ieee.org>)，以此作為取得 IEEE 和其他專業來源之智慧電網資訊的綜合入口。該網站專為製造商、政策制定者、教育工作者、學者、政府、工程師、電腦科學家、研究人員，電力 and 能源領域的其他利害關係人，以及 IT 和通訊產業所設計。IEEE 智慧電網專責小組主席 Wanda Reder 在 IEEE 在網站推出時表示：「需要電力和能源、通訊、資訊科技 (IT) 等產業，以及政府和學術界的齊心貢獻，才能確保全球成功實作智慧電網。」此外，IEEE 電力與能源協會 (PES) 一直積極參與各種層次的智慧電網活動。

另一個團體 GridWise Alliance，這是擁有 125 位成員組織的技術中立智慧電網聯盟，他們已與 IEEE PES 簽署協議，雙方將共同致力解決彼此都感興趣的智慧電網相關問題。

GridWise Alliance 的總裁，Katherine Hamilton 表示：「合作協議可以加強我們聯盟工程基礎。」

GridWise Alliance 還協助發起智慧電網消費者聯盟，重點放在教導消費者如何從智慧電網受惠 (參閱「何謂智慧電網？」)。

在消費者方面，智慧電網消費者聯盟的重點是推動消費者採用智慧電網。早期成員包括 GE、IBM、GridWise Alliance、Magnolia/Best Buy、Ember Corp.，以及美國能源部再生性能源實驗室等等。

有鑑於產業規模達 1650 億美元，消費性電子產品協會 (CEA) 特別成

立智慧電網工作小組，以針對會影響消費性電子產品的智慧電網專案和政策進行資訊交流。CEA 表示，他們將同時為 SGIP 和 NIST 提供意見。

家電製造商協會 (AHAM) 獲得 EnerNex 顧問公司的協助 (該公司也協助 NIST 管理智慧電網標準)，以針對所有家庭網路技術起草詳細的評論報告。AHAM 表示，希望他們的評估報告有助於決定，該如何就智慧電網問題向成員提出建議。

IEEE 已預定在 2010 年 10 月 4 至 6 日，於位在美國馬里蘭州 Gaithersburg 的 NIST 辦公室舉行 SmartGridComm 2010。這次會議的重點放在智慧電網的通訊層面，來自學術界、產業界和

國家實驗室的研究人員將在會議期間進行簡報並交換意見。會議相關資訊請參閱 <http://www.ieee-smartgrid-comm.org>。



Compressive Sampling

© DIGITAL VISION

An Introduction To Compressive Sampling

[A sensing/sampling paradigm that goes against the common knowledge in data acquisition]

[Emmanuel J. Candès
and Michael B. Wakin]

Conventional approaches to sampling signals or images follow Shannon's celebrated theorem: the sampling rate must be at least twice the maximum frequency present in the signal (the so-called Nyquist rate). In fact, this principle underlies nearly all signal acquisition protocols used in consumer audio and visual electronics, medical imaging devices, radio receivers, and so on. (For some signals, such as images that are not naturally bandlimited, the sampling rate is dictated not by the Shannon theorem but by the desired temporal or spatial resolution. However, it is common in such systems to use an antialiasing low-pass filter to bandlimit the signal before sampling, and so the Shannon theorem plays an implicit role.) In the field of data conversion, for example, standard analog-to-digital converter (ADC) technology implements the usual quantized Shannon representation: the signal is uniformly sampled at or above the Nyquist rate.

Digital Object Identifier 10.1109/MSP.2007.914731



Compressive Sampling

© DIGITAL VISION

壓縮取樣導論

【一種異于傳統資料獲取概念的感測/採樣範例】

Emmanuel J. Candès
Michael B. Wakin

傳

傳統訊號或圖像採樣方法都遵循著名的香農定理：採樣速率至少必須是受測訊號最高頻率之兩倍（奈奎斯特率）。實際上，該定理成為所有訊號採集協定之基礎，應用於家用影音電子產品，醫學影像設備，無線電接收設備中（對於一些訊號，例如並非本來就頻寬有限之圖像，採樣速率並不受制於香農定理，而受到期許的時間或空間解析度限制。然而，該系統常採用一個反混疊的低通濾波器在採樣之前限制訊號頻帶，這樣香農定理就扮演一個隱形的角色）。在資料轉換領域，例如，標準的類比/數位轉換器（ADC）技術則實現了香農表示方法之一般量化：在不小於奈奎斯特率時對訊號進行均勻採樣。

數位物件識別碼 10.1109/MSP.2007.914731

This article surveys the theory of compressive sampling, also known as compressed sensing or CS, a novel sensing/sampling paradigm that goes against the common wisdom in data acquisition. CS theory asserts that one can recover certain signals and images from far fewer samples or measurements than traditional methods use. To make this possible, CS relies on two principles: *sparsity*, which pertains to the signals of interest, and *incoherence*, which pertains to the sensing modality.

- Sparsity expresses the idea that the “information rate” of a continuous time signal may be much smaller than suggested by its bandwidth, or that a discrete-time signal depends on a number of degrees of freedom which is comparably much smaller than its (finite) length. More precisely, CS exploits the fact that many natural signals are sparse or compressible in the sense that they have concise representations when expressed in the proper basis Ψ .

- Incoherence extends the duality between time and frequency and expresses the idea that objects having a sparse representation in Ψ must be spread out in the domain in which they are acquired, just as a Dirac or a spike in the time domain is spread out in the frequency domain. Put differently, incoherence says that unlike the signal of interest, the sampling/sensing waveforms have an extremely dense representation in Ψ .

The crucial observation is that one can design efficient sensing or sampling protocols that capture the useful information content embedded in a sparse signal and condense it into a small amount of data. These protocols are nonadaptive and simply require correlating the signal with a small number of fixed waveforms that are incoherent with the sparsifying basis. What is most remarkable about these sampling protocols is that they allow a sensor to very efficiently capture the information in a sparse signal without trying to comprehend that signal. Further, there is a way to use numerical optimization to reconstruct the full-length signal from the small amount of collected data. In other words, CS is a very simple and efficient signal acquisition protocol which samples—in a signal independent fashion—at a low rate and later uses computational power for reconstruction from what appears to be an incomplete set of measurements.

Our intent in this article is to overview the basic CS theory that emerged in the works [1]–[3], present the key mathematical ideas underlying this theory, and survey a couple of important results in the field. Our goal is to explain CS as plainly as possible, and so our article is mainly of a tutorial nature. One of the charms of this theory is that it draws from various subdisciplines within the applied mathematical sciences, most notably probability theory. In this review, we have decided to highlight this aspect and especially the fact that randomness can—perhaps surprisingly—lead to very effective

sensing mechanisms. We will also discuss significant implications, explain why CS is a concrete protocol for sensing and compressing data simultaneously (thus the name), and conclude our tour by reviewing important applications.

THE SENSING PROBLEM

In this article, we discuss sensing mechanisms in which information about a signal $f(t)$ is obtained by linear functionals recording the values

$$y_k = \langle f, \varphi_k \rangle, \quad k = 1, \dots, m. \quad (1)$$

That is, we simply correlate the object we wish to acquire with the waveforms $\varphi_k(t)$. This is a standard setup. If the sensing waveforms are Dirac delta functions (spikes), for

example, then y is a vector of sampled values of f in the time or space domain. If the sensing waveforms are indicator functions of pixels, then y is the image data typically collected by sensors in a digital camera. If the sensing waveforms are sinusoids, then y is a vector of Fourier coef-

ficients; this is the sensing modality used in magnetic resonance imaging (MRI). Other examples abound.

Although one could develop a CS theory of continuous time/space signals, we restrict our attention to discrete signals $f \in \mathbb{R}^n$. The reason is essentially twofold: first, this is conceptually simpler and second, the available discrete CS theory is far more developed (yet clearly paves the way for a continuous theory—see also “Applications”). Having said this, we are then interested in *undersampled* situations in which the number m of available measurements is much smaller than the dimension n of the signal f . Such problems are extremely common for a variety of reasons. For instance, the number of sensors may be limited. Or the measurements may be extremely expensive as in certain imaging processes via neutron scattering. Or the sensing process may be slow so that one can only measure the object a few times as in MRI. And so on.

These circumstances raise important questions. Is accurate reconstruction possible from $m \ll n$ measurements only? Is it possible to design $m \ll n$ sensing waveforms to capture almost all the information about f ? And how can one approximate f from this information? Admittedly, this state of affairs looks rather daunting, as one would need to solve an underdetermined linear system of equations. Letting A denote the $m \times n$ sensing matrix with the vectors $\varphi_1^*, \dots, \varphi_m^*$ as rows (a^* is the complex transpose of a), the process of recovering $f \in \mathbb{R}^n$ from $y = Af \in \mathbb{R}^m$ is ill-posed in general when $m < n$: there are infinitely many candidate signals \tilde{f} for which $A\tilde{f} = y$. But one could perhaps imagine a way out by relying on realistic models of objects f which naturally exist. The Shannon

CS THEORY ASSERTS THAT ONE CAN RECOVER CERTAIN SIGNALS AND IMAGES FROM FAR FEWER SAMPLES OR MEASUREMENTS THAN TRADITIONAL METHODS USE.

本文分析了壓縮取樣理論，也稱之為壓縮感測或 CS，一個新的感測/採樣範例，與資料獲取中傳統概念不同。CS 理論強調一個確定訊號和圖像可以從比傳統方法更少的樣本或量測中重構。為了使之實現，CS 依賴兩個原理：稀疏性，與感興趣訊號之性質有關，非相干性，與感測形式有關。

■ 稀疏性表達了這樣的思想，一個連續時間訊號之「資訊速率」可能比由頻寬所決定的低很多，或者，一個離散時間訊號所依賴的自由度數目遠遠低於其（有限）長度。更準確地說，CS 利用了這一事實，即許多自然訊號在某種意義下是稀疏的或可壓縮的，因此在使用合適的基 ψ 表達時，其具有簡潔的表示形式。

■ 非相干性擴展時頻對偶性，表達這樣的思想，正如時間域 Dirac 或尖峰訊號可以在頻域展開那樣，在基 ψ 下具有稀疏表示的訊號一定可以獲得它們的某個域展開。但不同之處在於，非相干性表明，與感興趣訊號不同，用於採樣/感測的波形在基 ψ 上具有非常密集表示形式。至關重要的是人們可以設計出有效的感測或採樣協定，來捕捉嵌入稀疏訊號之有用資訊內容，從而將其壓縮成少量數據。這些協定是非自我調整的，僅僅需要將少量的固定波形和原始訊號關聯，而這些固定波形與為訊號提供簡潔表示的基為非相干。這些採樣協定最不同尋常之處是允許感測器非常有效地捕捉一個稀疏訊號中的資訊，但並未試圖理解該訊號。進一步而言，使用數值最佳化方法就可以從少量採集的資料中重構出完整長度之訊號。換而言之，CS 是一個非常簡單有效的訊號採集協定，以較低的採樣速率和獨立於訊號的採樣方式採樣訊號，然後又透過強大的計算能力從看上去不完整的量測集中重構原訊號。

我們在本文中概述文獻 [1]-[3] 中出現的基本 CS 理論，介紹這些理論中所包含之關鍵數學思想，分析該領域的一些重要成果。我們的目標是盡可能清晰地解釋 CS，所以本文主要是指導性質的。該理論的魅力之一是其源於數學學科內的不同分支，最明顯的範例就是機率論。在此綜述中，我們決定強調這個部分，尤其是隨機性可以一一也許是驚人地一一導向非常有效的量測機制這個事實。我們還將討論其重要意義，解釋 CS 為什麼是同時感測和壓縮資料（因而命名）的明確協定，最後以討論重要應用作為結束。

感測問題

本文中，我們將探討感測機制，其中由線性泛函得到的關於訊號 $f(t)$ 的資訊，其值記為：

$$y_k = \langle f, \varphi_k \rangle, \quad k = 1, \dots, m \quad (1)$$

也就是說，我們只簡單地將所期望得到的物件和波形 $\varphi_k(t)$ 關聯。這是一個標準架構。例如，如果感測波形是 Dirac 脈衝函數（尖峰訊號），那麼 y 就是 f 在頻域或時域上一個採樣值向量。如果感測波形是像素的指標函數，那麼 y 是由數位相機中的感測器特別採集的圖像資料。如果感測波形是正弦曲線，那麼 y 是傅立葉係數向量，這種感測模式被用於核磁共振影像 (MRI)。其它的範例亦很多。

儘管時間/空間連續訊號之 CS 理論已有所發展，但我們仍將注意力集中於離散訊號 $f \in \mathbb{R}^n$ 上。原因有兩部分：第一，概念上更簡單，第二，現有的離散 CS 理論還需更多發展（但已為

連續理論清晰地鋪平道路——見「應用」部分）。儘管如此，我們對欠採樣情況感興趣，此時可用的量測數目 m 遠少於訊號 f 的維數 n 。由於各種原因，該問題十分常見。例如，感測器數目有限。或者某些擷取影像過程借助中子散射，因而量測十分昂貴。或者感測過程可能太緩慢，所以量測目標需要一段時間，如 MRI。等等。

這些狀況產生了一些重要問題。根據 $m \ll n$ 次量測能精確重構訊號嗎？能設計出 $m \ll n$ 個感測波形來捕捉絕大部分關於 f 的資訊嗎？根據該資訊如何得到 f 之逼近？毫無疑問，這情形看起來都是比較棘手的事情，其中一個問題就是欠定線性方程組之求解。令 A 表示一個行向量為 $\varphi_1^*, \dots, \varphi_m^*$ ， $m \times n$ 大小的感測矩陣（ a^* 是 a 的複數轉置），當 $m < n$ 時，從 $y = Af \in \mathbb{R}^m$ 中恢復 $f \in \mathbb{R}^n$ 的過程通常不是適定的：這時對於 $A\tilde{f} = y$ ，有無窮多組解 \tilde{f} 。但是也許可以設想一種依賴於本來就存在的目標 f 之實際模型作為解決問題的方法。香農定理告訴我們，如果 $f(t)$ 實際上具有低頻寬，那麼少量的（均勻）樣本將足以將其恢復。正如本文其餘部分所述，實際上對於更多類型之訊號模型，訊號恢復亦有可能實現。

**CS 理論強調可以從比傳統方法
少得多的樣本或量測中
重構確定的訊號或圖像**

theory tells us that, if $f(t)$ actually has very low bandwidth, then a small number of (uniform) samples will suffice for recovery. As we will see in the remainder of this article, signal recovery can actually be made possible for a much broader class of signal models.

INCOHERENCE AND THE SENSING OF SPARSE SIGNALS

This section presents the two fundamental premises underlying CS: sparsity and incoherence.

SPARSITY

Many natural signals have concise representations when expressed in a convenient basis. Consider, for example, the image in Figure 1(a) and its wavelet transform in (b). Although nearly all the image pixels have nonzero values, the wavelet coefficients offer a concise summary: most coefficients are small, and the relatively few large coefficients capture most of the information.

Mathematically speaking, we have a vector $f \in \mathbb{R}^n$ (such as the n -pixel image in Figure 1) which we expand in an orthonormal basis (such as a wavelet basis) $\Psi = [\psi_1 \psi_2 \cdots \psi_n]$ as follows:

$$f(t) = \sum_{i=1}^n x_i \psi_i(t), \quad (2)$$

where x is the coefficient sequence of f , $x_i = \langle f, \psi_i \rangle$. It will be convenient to express f as Ψx (where Ψ is the $n \times n$ matrix with ψ_1, \dots, ψ_n as columns). The implication of sparsity is now clear: when a signal has a sparse expansion, one can discard the small coefficients without much perceptual loss. Formally, consider $f_S(t)$ obtained by keeping only the terms corresponding to the S largest values of (x_i) in the expansion (2). By definition, $f_S := \Psi x_S$, where here and below, x_S is the vector of coefficients (x_i) with all but the largest S set to zero. This vector is sparse in a strict sense since all but a few of its entries are zero; we will call S -sparse such objects with at most S nonzero entries. Since Ψ is an orthonormal basis (or “orthobasis”), we have $\|f - f_S\|_{\ell_2} = \|x - x_S\|_{\ell_2}$, and if x is sparse or *compressible* in the sense that the sorted magnitudes of the (x_i) decay quickly, then x is well approximated by x_S and, therefore, the error $\|f - f_S\|_{\ell_2}$ is small. In plain terms, one can “throw away” a large fraction of the coefficients without much loss. Figure 1(c) shows an example where the perceptual loss is hardly noticeable from a megapixel image to its approximation obtained by throwing away 97.5% of the coefficients.

This principle is, of course, what underlies most modern lossy coders such as JPEG-2000 [4] and many

others, since a simple method for data compression would be to compute x from f and then (adaptively) encode the locations and values of the S significant coefficients. Such a process requires knowledge of all the n coefficients x , as the locations of the significant pieces of information may not be known in advance (they are signal dependent); in our example, they tend to be clustered around edges in the image. More generally, sparsity is a fundamental modeling tool which permits efficient fundamental signal processing; e.g., accurate statistical estimation and classification, efficient data compression, and so on. This article is about a more surprising and far-reaching implication, however, which is that sparsity has significant bearings on the acquisition process itself. Sparsity determines how efficiently one can acquire signals *nonadaptively*.

INCOHERENT SAMPLING

Suppose we are given a pair (Φ, Ψ) of orthobases of \mathbb{R}^n . The first basis Φ is used for sensing the object f as in (1) and the second is used to represent f . The restriction to pairs of orthobases is not essential and will merely simplify our treatment.

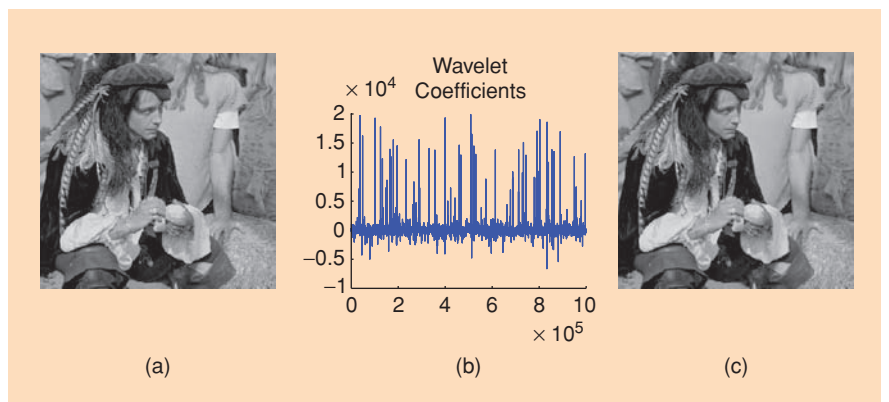
DEFINITION 1

The coherence between the sensing basis Φ and the representation basis Ψ is

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max_{1 \leq k, j \leq n} |\langle \varphi_k, \psi_j \rangle|. \quad (3)$$

In plain English, the coherence measures the largest correlation between any two elements of Φ and Ψ ; see also [5]. If Φ and Ψ contain correlated elements, the coherence is large. Otherwise, it is small. As for how large and how small, it follows from linear algebra that $\mu(\Phi, \Psi) \in [1, \sqrt{n}]$.

Compressive sampling is mainly concerned with low coherence pairs, and we now give examples of such pairs. In our first example, Φ is the canonical or spike basis $\varphi_k(t) = \delta(t - k)$ and



[FIG1] (a) Original megapixel image with pixel values in the range [0,255] and (b) its wavelet transform coefficients (arranged in random order for enhanced visibility). Relatively few wavelet coefficients capture most of the signal energy; many such images are highly compressible. (c) The reconstruction obtained by zeroing out all the coefficients in the wavelet expansion but the 25,000 largest (pixel values are thresholded to the range [0,255]). The difference with the original picture is hardly noticeable. As we describe in “Undersampling and Sparse Signal Recovery,” this image can be perfectly recovered from just 96,000 incoherent measurements.

非相干性和稀疏訊號的感測

該章節提出了 CS 的兩個基本前提：稀疏性和非相干性。

稀疏性

許多自然訊號在一些傳統基下具有簡潔表示。例如，圖 1(a) 中的圖像及其小波轉換 (b)。雖然幾乎所有的圖像像素都是非零值，但小波係數提出了一個簡潔的概括：大部分係數非常小，相較而言少量的大係數捕捉到大部分資訊。

從數學角度來講，我們將一個向量 $f \in \mathbb{R}^n$ (例如圖 1 中具有 n 個像素點的圖像) 在正交基 $\Psi = [\psi_1 \psi_2 \dots \psi_n]$ (例如小波基) 下展開：

$$f(t) = \sum_{i=1}^n x_i \psi_i(t) \quad (2)$$

這裡 x 是 f 的係數序列， $x_i = \langle f, \psi_i \rangle$ 。為了方便起見，將 f 表示為 Ψx (這裡 Ψ 是行向量 ψ_1, \dots, ψ_n 組成的 $n \times n$ 大小的矩陣)。稀疏性的含義現在非常明確：當一個訊號有稀疏展開式時，丟棄小係數並沒有造成大量感知損失。就形式上來說， $f_S(t)$ 由僅保留展開式(2)中對應於 S 個 (x_i) 最大值的項而獲取。定義 $f_S := \Psi x_S$ ，此後， x_S 表示係數 (x_i) 向量，但除 S 個大係數集合外其餘都置零。該向量從嚴格意義上來說是稀疏的，因為除了少數非零元素外，其餘均為零；我們稱之為 S -稀疏的，這樣目標至多有 S 個非零項。由於 ψ 是一個標準正交基 (或「正交基」)，我們有 $\|f - f_S\|_{\ell_2} = \|x - x_S\|_{\ell_2}$ ，如果 x 是稀疏或可壓縮的，這就意味著 (x_i) 按值排序快速衰減，那麼 x 就能用 x_S 很好地逼近，因此誤差 $\|f - f_S\|_{\ell_2}$ 較小。簡單地說，就是「拋棄」很大部分的係數而沒有造成較大損失。圖 1(c) 提出了這樣的範例：將一幅百萬像素的圖像在丟掉了 97.5% 的係數後得到其逼近，可視覺損失幾乎難以察覺。

該定理構成了大部分現代有損編碼的基礎，例如，JPEG-2000 [4] 或者許多其它編碼器，因為資料壓縮的簡單方法是從 f 中計算 x ，然後 (自我調整地) 編碼 S 個重要係數的數值和位置。這樣的過程需要知道所有的 n 個係數 x ，因為資訊的重要片段所在的位置可能事先未知 (取決於訊號)；在我們的範例中，

它們有著沿著圖像邊緣叢集的趨勢。一般而言，稀疏性是一個基本的建模工具，使有效的基本訊號處理成為可能；例如，準確的統計估計和分類，有效的資料壓縮，等等。但是，本文研究的是比以上更異乎尋常和更具深遠意義之題目，但是，稀疏性在自身的採集過程中具有重要影響。稀疏性決定了如何有效地對訊號進行非自我調整的採集。

非相干採樣

假設我們給定一對 \mathbb{R}^n 內的正交基 (Φ, Ψ)。用第一個基來感測公式 (1) 之目標 f ，然後用第二個基來表徵 f 。對該正交基對的限制並非必需，僅僅為了簡化我們的處理。

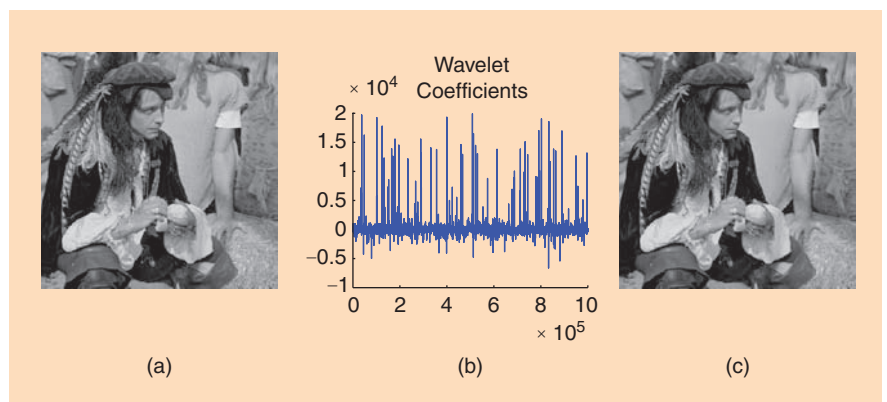
定義 1

感測基 (系統) Φ 和表示基 (系統) Ψ 之間的相干性定義如下：

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max_{1 \leq k, j \leq n} |\langle \varphi_k, \psi_j \rangle| \quad (3)$$

簡單地講，相干性度量的是兩個矩陣 Φ 和 Ψ 的元素之間的最大相關性；見文獻 [5]。如果 Φ 和 Ψ 包含相關的元素，則相關性會很大；否則就會很小。相關係數取值範圍為 $\mu(\Phi, \Psi) \in [1, \sqrt{n}]$ 。

壓縮取樣主要研究具有低相干性的兩個系統，我們現在提出一些範例。第一個範例， Φ 是正則或尖峰基 $\varphi_k(t) = \delta(t - k)$ ， Ψ 為傅立葉基， $\psi_j(t) = n^{-1/2} e^{i2\pi j t/n}$ 。因為 Φ 是感測矩陣，對應於時域或空域中經典的採樣方案。時頻矩陣服從 $\mu(\Phi, \Psi) = 1$ ，因此我們得到最大程度的非相干性。進一步講，尖峰訊號和正弦訊號不僅在一



【圖 1】 (a) 原始的百萬像素圖像，其像素值範圍為 [0-255]，(b) 小波轉換係數 (為了加強視覺效果採用隨機排列順序)。相當少的小波係數捕捉到大部分的訊號能量；許多這類圖像皆可高度壓縮。(c) 除了 25,000 個最大值 (像素值被限定到 [0-255] 範圍) 外，所有其他的小波展開置零來獲得重構。與原始圖像的區別難以分辨。正如我們在「欠採樣和稀疏訊號恢復」中所描述的那樣，該圖像可以從 96,000 個非相干量測中完整恢復。

Ψ is the Fourier basis, $\psi_j(t) = n^{-1/2} e^{i 2\pi jt/n}$. Since Φ is the sensing matrix, this corresponds to the classical sampling scheme in time or space. The time-frequency pair obeys $\mu(\Phi, \Psi) = 1$ and, therefore, we have *maximal incoherence*. Further, spikes and sinusoids are maximally incoherent not just in one dimension but in any dimension, (in two dimensions, three dimensions, etc.)

Our second example takes wavelets bases for Ψ and noiselets [6] for Φ . The coherence between noiselets and Haar wavelets is $\sqrt{2}$ and that between noiselets and Daubechies D4 and D8 wavelets is, respectively, about 2.2 and 2.9 across a wide range of sample sizes n . This extends to higher dimensions as well. (Noiselets are also maximally incoherent with spikes and incoherent with the Fourier basis.) Our interest in noiselets comes from the fact that 1) they are incoherent with systems providing sparse representations of image data and other types of data, and 2) they come with very fast algorithms; the noiselet transform runs in $O(n)$ time, and just like the Fourier transform, the noiselet matrix does not need to be stored to be applied to a vector. This is of crucial practical importance for numerically efficient CS implementations.

Finally, random matrices are largely incoherent with any fixed basis Ψ . Select an orthobasis Φ uniformly at random, which can be done by orthonormalizing n vectors sampled independently and uniformly on the unit sphere. Then with high probability, the coherence between Φ and Ψ is about $\sqrt{2} \log n$. By extension, random waveforms ($\varphi_k(t)$) with independent identically distributed (i.i.d.) entries, e.g., Gaussian or ± 1 binary entries, will also exhibit a very low coherence with any fixed representation Ψ . Note the rather strange implication here; if sensing with incoherent systems is good, then efficient mechanisms ought to acquire correlations with random waveforms, e.g., white noise!

UNDERSAMPLING AND SPARSE SIGNAL RECOVERY

Ideally, we would like to measure all the n coefficients of f , but we only get to observe a subset of these and collect the data

$$y_k = \langle f, \varphi_k \rangle, \quad k \in M, \quad (4)$$

where $M \subset \{1, \dots, n\}$ is a subset of cardinality $m < n$. With this information, we decide to recover the signal by ℓ_1 -norm minimization; the proposed reconstruction f^* is given by $f^* = \Psi x^*$, where x^* is the solution to the convex optimization program ($\|x\|_{\ell_1} := \sum_i |x_i|$)

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_{\ell_1} \quad \text{subject to} \quad y_k = \langle \varphi_k, \Psi \tilde{x} \rangle, \quad \forall k \in M. \quad (5)$$

That is, among all objects $\tilde{f} = \Psi \tilde{x}$ consistent with the data, we pick that whose coefficient sequence has minimal ℓ_1 norm. (As is well known, minimizing ℓ_1 subject to linear equality con-

straints can easily be recast as a linear program making available a host of ever more efficient solution algorithms.)

The use of the ℓ_1 norm as a sparsity-promoting function traces back several decades. A leading early application was reflection seismology, in which a sparse reflection function (indicating meaningful changes between subsurface layers) was sought from bandlimited data [7], [8]. However, ℓ_1 -minimization is not the only way to

recover sparse solutions; other methods, such as greedy algorithms [9], have also been proposed.

Our first result asserts that when f is sufficiently sparse, the recovery via ℓ_1 -minimization is provably exact.

THEOREM 1 [10]

Fix $f \in \mathbb{R}^n$ and suppose that the coefficient sequence x of f in the basis Ψ is S -sparse. Select m measurements in the Φ domain uniformly at random. Then if

$$m \geq C \cdot \mu^2(\Phi, \Psi) \cdot S \cdot \log n \quad (6)$$

for some positive constant C , the solution to (5) is exact with overwhelming probability. (It is shown that the probability of success exceeds $1 - \delta$ if $m \geq C \cdot \mu^2(\Phi, \Psi) \cdot S \cdot \log(n/\delta)$. In addition, the result is only guaranteed for nearly all sign sequences x with a fixed support, see [10] for details.)

We wish to make three comments:

- 1) The role of the coherence is completely transparent; the smaller the coherence, the fewer samples are needed, hence our emphasis on low coherence systems in the previous section.
- 2) One suffers no information loss by measuring just about any set of m coefficients which may be far less than the signal size apparently demands. If $\mu(\Phi, \Psi)$ is equal or close to one, then on the order of $S \log n$ samples suffice instead of n .
- 3) The signal f can be exactly recovered from our condensed data set by minimizing a convex functional which does not assume any knowledge about the number of nonzero coordinates of x , their locations, or their amplitudes which we assume are all completely unknown a priori. We just run the algorithm and if the signal happens to be sufficiently sparse, exact recovery occurs.

The theorem indeed suggests a very concrete acquisition protocol: sample nonadaptively in an incoherent domain and invoke linear programming after the acquisition step. Following this protocol would essentially acquire the signal in a compressed form. All that is needed is a decoder to “decompress” this data; this is the role of ℓ_1 minimization.

In truth, this random incoherent sampling theorem extends an earlier result about the sampling of spectrally sparse signals [1], which showed that randomness 1) can be a very effective

MANY NATURAL SIGNALS ARE SPARSE OR COMPRESSIBLE IN THE SENSE THAT THEY HAVE CONCISE REPRESENTATIONS WHEN EXPRESSED IN THE PROPER BASIS.

維而且在任何維都具有最大的非相干性（在二維，三維空間等等）。

第二個範例中 Ψ 採用小波基， Φ 是 noiselets [6]。Noiselets 和 Haar 小波基間的相干係數為 $\sqrt{2}$ ，noiselets 超過了樣本容量為 n 的範圍，與 Daubechies D4 和 D8 小波基之間的相干係數分別是 2.2 和 2.9。這也可以擴展到高維情況（noiselets 也和尖峰訊號及傅立葉基最大不相關）。人們對 noiselets 感興趣是基於如下兩個事實：1) 它們與提供圖像資料和其它類型資料稀疏表示的系統非相干，2) 具有快速演算法，noiselet 轉換的時間複雜度為 $O(n)$ ，與傅立葉轉換一樣，noiselet 矩陣應用於向量時不需要存儲。對數值高效的 CS 實現來說，實用性至關重要。

最後一個範例，隨機矩陣可以是任何固定基 Ψ 之間具有極大的非相干。透過 n 個向量在單位球面上獨立均勻地採樣並做規範正交化，可以隨機均勻地選擇正交基 Φ 。此時， Φ 和 Ψ 間的相干性在很大機率上為 $\sqrt{2 \log n}$ 。推而廣之，各項服從獨立同分佈 (i.i.d.) 的隨機波形 ($\varphi_k(t)$)，例如高斯分佈或者 ± 1 二元項，也表現出和任何固定基 Ψ 間具有較小的相干性。注意在這裡有個非常獨特的含意：如果使用非相干系統感測良好，那麼有效機制應該獲取到同隨機波形間的關聯，例如白色雜訊。

欠採樣和稀疏訊號恢復

理想狀態下，我們希望對 f 的 n 個係數全部進行感測，但我們僅能感測其中的子集，並採集資料：

$$y_k = \langle f, \varphi_k \rangle, \quad k \in M \quad (4)$$

這裡 $M \subset \{1, \dots, n\}$ 是基數 $m < n$ 的一個子集。根據此資訊，我們決定由 ℓ_1 範數最小化來恢復訊號；所提出的重構 f 由 $f^* = \Psi x^*$ 提出，這裡 x^* 是凸型最佳化問題 ($\|x\|_{\ell_1} := \sum_i |x_i|$) 之解

$$\min \| \tilde{x} \|_{\ell_1} \text{ subject to } y_k = \langle \varphi_k, \Psi \tilde{x} \rangle, \quad \forall k \in M \quad (5)$$

即，在所有與資料均非相干的目標 $\tilde{f} = \Psi \tilde{x}$ 中，我們選擇那些具有最小 ℓ_1 範數的係數序列（眾所周知，服從線性等式約束條件的 ℓ_1 最小化很容易轉化為一個線性規劃，可找到更高效的求解演算法）。

許多自然訊號在某種意義下是稀疏的或可壓縮的，在用合適的基表達時，它們具有簡潔的表示形式

ℓ_1 範數用作稀疏提升函數可以追溯到幾十年前。早期的主要應用是反射地震學，從頻寬有限的資料中尋找一個稀疏反射函數（表明次表面層間的重要變化）[7] [8]。然而 ℓ_1 極小化並不是恢復稀疏解的唯一方法；也有其它方法，例如已經提出的貪婪演算法[9]。我們最初成果表明，當 f 足夠稀疏時，可以證明，透過 ℓ_1 極小化可以準確恢復訊號。

定理1 [10]

給定 $f \in \mathbb{R}^n$ 且假設 f 的係數序列 x 在基 Ψ 中是 S -稀疏的。均勻地在 Φ 域均勻隨機選擇 m 個量測。那麼對於某一正常數 C 有

$$m \geq C \cdot \mu^2(\Phi, \Psi) \cdot S \cdot \log n \quad (6)$$

公式(5)的解可以充分保證是準確的（這提出了如果 $m \geq C \cdot \mu^2(\Phi, \Psi) \cdot S \cdot \log(n/\delta)$ 成功超過 $1 - \delta$ 的機率，另外，該結果僅保證是針對幾乎所有具有固定支撐的符號序列 x ，詳見文獻 [10]）。

我們提出下面三點解釋：

- 1) 相干性的作用非常明顯；相干性越小，所需的樣本越少，因此我們前面的章節重點強調低相干性的系統。
- 2) 僅量測任一 m 係數集合時，都無資訊損失，而此係數可能遠少於訊號規模的明確要求。如果 $\mu(\Phi, \Psi)$ 等於或接近1，那麼只需大約 $S \log n$ 個樣本就足夠，而不是 n 個。
- 3) 在事先未知任何關於 x 非零座標個數、位置，以及幅值條件下，透過凸函數極小化從壓縮資料集中準確恢復訊號 f 。我們僅僅運行演算法即可。如果訊號剛好足夠稀疏，就可以精確恢復。

該定理的確提出了一個非常具體的採集協定：在一個非相干域進行非自我調整採樣，採集步驟後執行線性規劃。按照此協定，從本質上獲取了壓縮形式的訊號。再需要一個解碼器來「解壓縮」該資料，這就是 ℓ_1 極小化所起的作用。

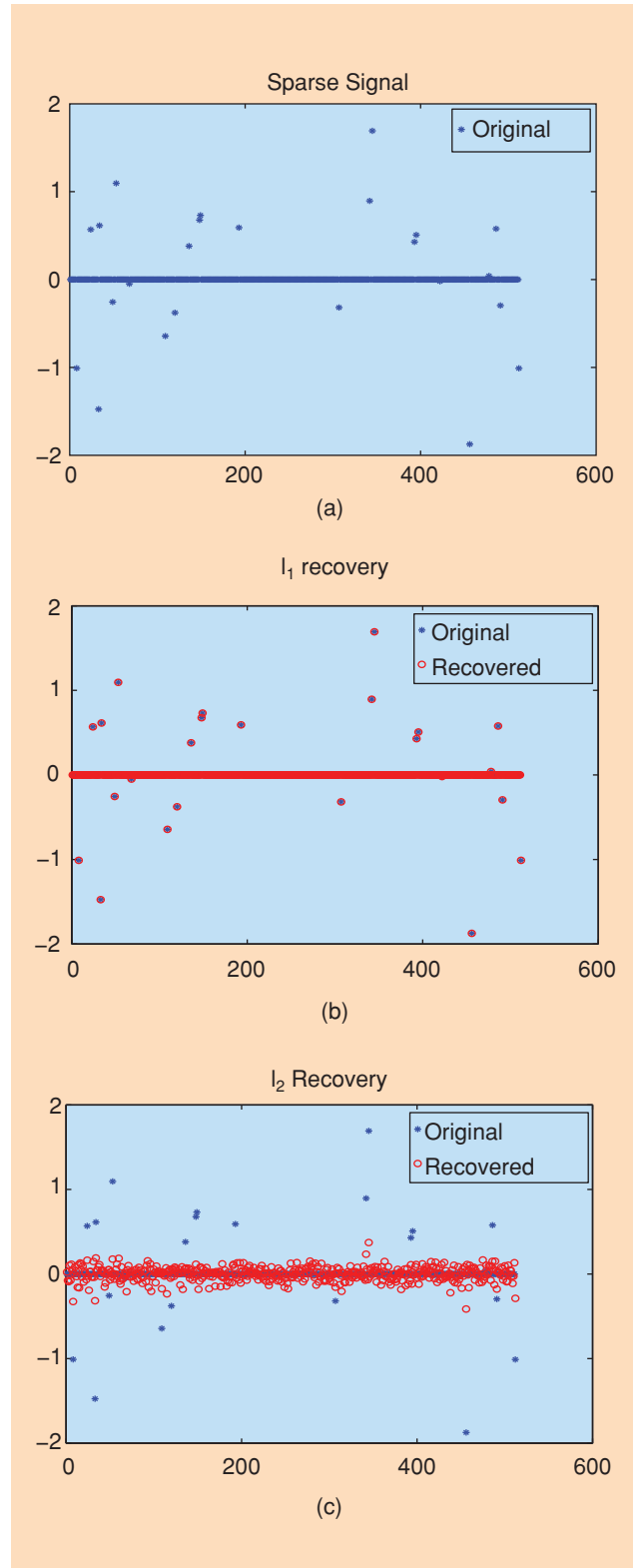
事實上，隨機非相干採樣理論擴展了早期成果中關於譜稀疏訊號的採樣 [1]，這表明隨機性 1) 可以成爲一個高效的感測機制；2) 應服從嚴格的證明，也許因此引發了我們已經見證和在今天將繼續見證的許多 CS 發展。假設我們採樣超頻寬卻譜稀疏的訊號，其形式如下： $f(t) = \sum_{j=0}^{n-1} x_j e^{j2\pi i t/n}$, $t = 0, \dots, n-1$ ，這裡 n 非

sensing mechanism and 2) is amenable to rigorous proofs, and thus perhaps triggered the many CS developments we have witnessed and continue to witness today. Suppose that we are interested in sampling ultra-wideband but spectrally sparse signals of the form $f(t) = \sum_{j=0}^{n-1} x_j e^{i2\pi jt/n}$, $t = 0, \dots, n-1$, where n is very large but where the number of nonzero components x_j is less than or equal to S (which we should think of as comparably small). We do not know which frequencies are active nor do we know the amplitudes on this active set. Because the active set is not necessarily a subset of consecutive integers, the Nyquist/Shannon theory is mostly unhelpful (since one cannot restrict the bandwidth a priori, one may be led to believe that all n time samples are needed). In this special instance, Theorem 1 claims that one can reconstruct a signal with arbitrary and unknown frequency support of size S from on the order of $S \log n$ time samples, see [1]. What is more, these samples do not have to be carefully chosen; almost any sample set of this size will work. An illustrative example is provided in Figure 2. For other types of theoretical results in this direction using completely different ideas see [11]–[13].

It is now time to discuss the role played by probability in all of this. The key point is that to get useful and powerful results, one needs to resort to a probabilistic statement since one cannot hope for comparable results holding for all measurement sets of size m . Here is why. There are special sparse signals that vanish nearly everywhere in the Φ domain. In other words, one can find sparse signals f and very large subsets of size almost n (e.g., $n - S$) for which $y_k = \langle f, \phi_k \rangle = 0$ for all $k \in M$. The interested reader may want to check the example of the Dirac comb discussed in [14] and [1]. On the one hand, given such subsets, one would get to see a stream of zeros and no algorithm whatsoever would of course be able to reconstruct the signal. On the other hand, the theorem guarantees that the fraction of sets for which exact recovery does not occur is truly negligible (a large negative power of n). Thus, we only have to tolerate a probability of failure that is extremely small. For practical purposes, the probability of failure is zero provided that the sampling size is sufficiently large.

Interestingly, the study of special sparse signals discussed above also shows that one needs at least on the order of $\mu^2 \cdot S \cdot \log n$ samples as well. (We are well aware that there exist subsets of cardinality $2S$ in the time domain which can reconstruct any s -sparse signal in the frequency domain. Simply take $2s$ consecutive time points, see “What Is Comprehensive Sampling?” and [11] and [12], for example. But this is not what our claim is about. We want that most sets of a certain size provide exact reconstruction.) With fewer samples, the probability that information may be lost is just too high and reconstruction by any method, no matter how intractable, is impossible. In summary, when the coherence is one, say, we do not need more than $S \log n$ samples but we cannot do with fewer either.

We conclude this section with an incoherent sampling example, and consider the sparse image in Figure 1(c), which as we recall has only 25,000 nonzero wavelet coefficients. We then



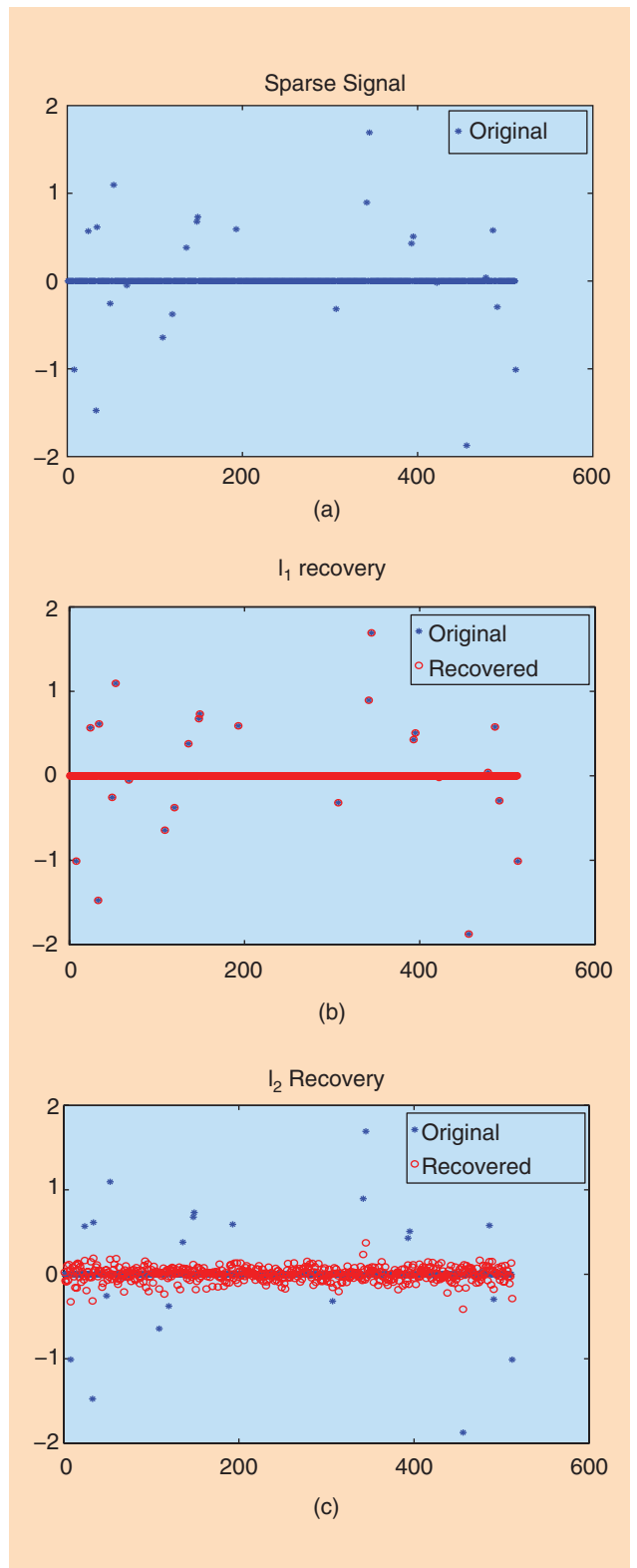
[FIG2] (a) A sparse real valued signal and (b) its reconstruction from 60 (complex valued) Fourier coefficients by ℓ_1 minimization. The reconstruction is exact. (c) The minimum energy reconstruction obtained by substituting the ℓ_1 norm with the ℓ_2 norm; ℓ_1 and ℓ_2 give wildly different answers. The ℓ_2 solution does not provide a reasonable approximation to the original signal.

常大但非零分量 x_i 的數目少於或等於 S （應該認為是非常小）。我們不知道哪些頻率是有效的，也不知道有效頻率集合上的幅度。因為有效集合不一定是一個連續整數的子集，奈奎斯特/香農定理毫無作用（因為不能限定先驗頻寬，甚至認為所有的 n 次採樣都必需）。在該特例中，定理 1 說明從大約 $S \log n$ 個樣本中就可以重構容量為 S 的任意和未知頻率支撐的訊號，見文獻[1]。而且，這些採樣無需仔細選擇，該容量的幾乎所有的樣本集合都起作用。圖 2 提出了解說性範例。對於在該方向採取完全不同之思想的其它類型理論研究成果，請見文獻 [11]-[13]。

下面該討論機率在所有這些理論裡扮演的角色。關鍵在於為了得到有效和有利的結果，需要訴諸一個機率表述，因為不能希望對所有 m 大小的量測集合都有類似的結果。這裡提出原因。一些特殊的稀疏訊號幾乎在 Φ 域的各處都為零。換句話說，對於 $y_k = \langle f, \varphi_k \rangle = 0$ ，所有的 $k \in M$ ，可以尋找稀疏訊號 f ，及規模幾乎為 n 的非常大的子集（例如 $n - S$ ）。有興趣的讀者可以檢核文獻 [14] 和 [1] 中討論的 Dirac 脈衝波形之範例。一方面，給定子集，人們只能看到一個零串流，當然沒有任何演算法可以重構訊號。另一方面，該理論確保了資料集中不能進行準確恢復的小部分實際上可忽略（一個 n 的較大負次幕）。這樣我們必須容許一個相當小的失敗機率。從實用角度來講，如果採樣次數足夠大，則失敗機率為零。

有趣的是，對上述討論的特定稀疏訊號的研究也說明至少需要大約 $\mu^2 \cdot S \cdot \log n$ 個樣本。（我們注意到在時域存在基數 $2S$ 的子集，可以重構任一頻域中的 S -稀疏訊號，簡單地採用 $2S$ 個連續時間點，詳見「什麼是綜合採樣」和文獻 [11] [12]。但這不是我們所要陳述的內容。我們希望大部分規模確定的集合能提供精確重構）。用少數樣本，訊號有可能損失的機率實在太高，用任何多麼複雜方法來重構都不可能。總而言之，當相干性是 1 時說明，我們所需要的樣本不必多於 $S \log n$ 個，但也不能再少。

我們以一個非相干性採樣的範例來總結本節內容，考慮圖 1(c) 中的稀疏訊號，我們印象中只有 25,000 個非零小波係數。然後，我們用 96,000 次非相干量測來採集資訊（這些量測的詳細資料詳見文獻 [10]），並求解公式 (5)。 l_1 極小化是完全恢復，也就是說， $\hat{f} = f$ 。此範例說明樣本數目僅大約為稀疏水準的四倍即可。許多學者發表了類似的成功經驗。事實上有一個已知的四比一實用規則，其中表明為了準確重構，每一個未知的非零項大約需要四個非相干樣本。



【圖 2】(a) 一個稀疏的實值訊號 (b) 採用 l_1 極小化從 60 個（複值）傅立葉係數中重構該訊號，重構是準確的 (c) 由範數 l_2 取代範數 l_1 獲得的能量最小化重構， l_1 和 l_2 提出了完全不同的答案。 l_2 的解不能提供對原訊號的合適逼近。

acquire information by taking 96,000 incoherent measurements (see [10] for the particulars of these measurements) and solve (5). The minimum- ℓ_1 recovery is *perfect*; that is, $f^* = f$. This example shows that a number of samples just about $4 \times$ the sparsity level suffices. Many researchers have reported on similar empirical successes. There is de facto a known four-to-one practical rule which says that for exact recovery, one needs about four incoherent samples per unknown nonzero term.

WHAT IS MOST REMARKABLE ABOUT THESE SAMPLING PROTOCOLS IS THAT THEY ALLOW A SENSOR TO VERY EFFICIENTLY CAPTURE THE INFORMATION IN A SPARSE SIGNAL WITHOUT TRYING TO COMPREHEND THAT SIGNAL.

ROBUST COMPRESSIVE SAMPLING

We have shown that one could recover sparse signals from just a few measurements but in order to be really powerful, CS needs to be able to deal with both nearly sparse signals and with noise. First, general objects of interest are not exactly sparse but approximately sparse. The issue here is whether or not it is possible to obtain accurate reconstructions of such objects from highly undersampled measurements. Second, in any real application measured data will invariably be corrupted by at least a small amount of noise as sensing devices do not have infinite precision. It is therefore imperative that CS be robust vis a vis such nonidealities. At the very least, small perturbations in the data should cause small perturbations in the reconstruction.

This section examines these two issues simultaneously. Before we begin, however, it will ease the exposition to consider the abstract problem of recovering a vector $x \in \mathbb{R}^n$ from data

$$y = Ax + z, \tag{7}$$

where A is an $m \times n$ “sensing matrix” giving us information about x , and z is a stochastic or deterministic unknown error term. The setup of the last section is of this form since with $f = \Psi x$ and $y = R\Phi f$ (R is the $m \times n$ matrix extracting the sampled coordinates in \mathcal{M}), one can write $y = Ax$, where $A = R\Phi\Psi$. Hence, one can work with the abstract model (7) bearing in mind that x may be the coefficient sequence of the object in a proper basis.

RESTRICTED ISOMETRIES

In this section, we introduce a key notion that has proved to be very useful to study the general robustness of CS; the so-called *restricted isometry property* (RIP) [15].

DEFINITION 2

For each integer $S = 1, 2, \dots$, define the isometry constant δ_S of a matrix A as the smallest number such that

$$(1 - \delta_S)\|x\|_{\ell_2}^2 \leq \|Ax\|_{\ell_2}^2 \leq (1 + \delta_S)\|x\|_{\ell_2}^2 \tag{8}$$

holds for all S -sparse vectors x .

We will loosely say that a matrix A obeys the RIP of order S if δ_S is not too close to one. When this property holds, A approxi-

mately preserves the Euclidean length of S -sparse signals, which in turn implies that S -sparse vectors cannot be in the null space of A . (This is useful as otherwise there would be no hope of reconstructing these vectors.) An equivalent description of the RIP is to say that all subsets of S columns taken from A are in fact nearly orthogonal (the columns of A cannot be exactly orthogonal since we have more columns than rows).

To see the connection between the RIP and CS, imagine we wish to acquire S -sparse signals with

A . Suppose that δ_{2S} is sufficiently less than one. This implies that all pairwise distances between S -sparse signals must be well preserved in the measurement space. That is, $(1 - \delta_{2S})\|x_1 - x_2\|_{\ell_2}^2 \leq \|Ax_1 - Ax_2\|_{\ell_2}^2 \leq (1 + \delta_{2S})\|x_1 - x_2\|_{\ell_2}^2$ holds for all S -sparse vectors x_1, x_2 . As demonstrated in the next section, this encouraging fact guarantees the existence of efficient and robust algorithms for discriminating S -sparse signals based on their compressive measurements.

GENERAL SIGNAL RECOVERY FROM UNDERSAMPLED DATA

If the RIP holds, then the following linear program gives an accurate reconstruction:

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_{\ell_1} \quad \text{subject to} \quad A\tilde{x} = y (= Ax). \tag{9}$$

THEOREM 2 [16]

Assume that $\delta_{2S} < \sqrt{2} - 1$. Then the solution x^* to (9) obeys

$$\|x^* - x\|_{\ell_2} \leq C_0 \cdot \|x - x_S\|_{\ell_1} / \sqrt{S} \quad \text{and} \tag{10}$$

$$\|x^* - x\|_{\ell_1} \leq C_0 \cdot \|x - x_S\|_{\ell_1}$$

for some constant C_0 , where x_S is the vector x with all but the largest S components set to 0. (As stated, this result is due to the first author [17] and yet unpublished, see also [16] and [18].)

The conclusions of Theorem 2 are stronger than those of Theorem 1. If x is S -sparse, then $x = x_S$ and, thus, the recovery is exact. But this new theorem deals with all signals. If x is not S -sparse, then (10) asserts that the quality of the recovered signal is as good as if one knew ahead of time the location of the S largest values of x and decided to measure those directly. In other words, the reconstruction is nearly as good as that provided by an oracle which, with full and perfect knowledge about x , extracts the S most significant pieces of information for us.

Another striking difference with our earlier result is that it is deterministic; it involves no probability. If we are fortunate enough to hold a sensing matrix A obeying the hypothesis of

穩健的壓縮取樣

我們已證明可以由少量量測來恢復稀疏訊號，但是爲了實際上更有成效起見，CS 需要既能夠處理近似的稀疏訊號又能夠處理雜訊。首先，一般而言，感興趣的目標不是精確稀疏但卻是近似稀疏的。這裡的問題是，是否能夠從高度欠採樣資料獲得該目標的準確重構。其次，在任何實際應用中，量測的資料總是被存在的少量雜訊所汙損，因爲感測設備不能無限度精確。因此 CS 在面對諸如非理想性時的穩健性非常有必要。至少在資料中的微小干擾有可能引起重構之微小干擾。

本小節中同時審查這兩個問題。在我們開始前，先考慮從以下資料中恢復向量 $x \in \mathbb{R}^n$ 的抽象問題：

$$y = Ax + z \quad (7)$$

這裡 A 是一個 $m \times n$ 「感測矩陣」，提供關於 x 的資訊， z 是一個隨機或確定的未知誤差項。從上節 $f = \Psi x$ 和 $y = R\Phi f$ (R 是在 M 上提取採樣座標的 $m \times n$ 矩陣)，可以寫爲 $y = Ax$ ，這裡令 $A = R\Phi\Psi$ 。因此，考慮到 x 可以是以在合適基下表示目標的係數序列，人們可以與抽象模型 (7) 一起使用。

有限等距

本節中，我們引入一個關鍵概念，對研究 CS 的一般穩健性非常有用；稱之爲有限等距性質(restricted isometries, RIP) [15]。

定義 2

對於每一整數 $S = 1, 2, \dots$ ，定義矩陣 A 的等距常數 δ_S ，當其最小時有：

$$(1 - \delta_S) = \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_S) \|x\|_2^2 \quad (8)$$

適用於所有 S -稀疏的向量 x 。

如果 δ_S 並不太接近於 1，寬鬆地講，矩陣 A 是服從 S 階 RIP。當該性質成立時， A 近似地保持 S -稀疏訊號的歐氏範數，這相應地意味著 S -稀疏向量不存在於 A 之零空間中（這非常有用否則不可能重構這些向量）。RIP 的等

價描述指出取自 A 的所有 S 列子集實際上幾乎是正交的（ A 的行向量不可能精確正交，因爲行大於列）。

爲了得到 RIP 和 CS 之間的關聯，我們希望透過 A 得到 S -稀疏訊號。假設比 1 小很多。這表示所有 S -稀疏訊號間成對的距離都保留在量測空間中。也就是 $(1 - \delta_{2S}) = \|x_1 - x_2\|_2^2 \leq \|Ax_1 - Ax_2\|_2^2 \leq (1 + \delta_{2S}) \|x_1 - x_2\|_2^2$ 對所有的 S -稀疏向量 x_1, x_2 成立。如下面章節所論證，這促進了如下事實，確保高效和穩健的演算法存在，從壓縮量測來區別 S -稀疏訊號。

從欠採樣資料恢復一般訊號

假如滿足 RIP，那麼則依線性規劃就提出精確重構：

$$\min_{\tilde{x} \in \mathbb{R}^n} \|\tilde{x}\|_{\ell_1} \quad \text{subject to} \quad A\tilde{x} = y (= Ax) \quad (9)$$

定理 2 [16]

假設 $\delta_{2S} < \sqrt{2} - 1$ 。那麼公式(9)的解 x^* 對於某一常數 C_0 而言，服從：

$$\|x^* - x\|_{\ell_2} \leq C_0 \cdot \|x - x_S\|_{\ell_1} / \sqrt{S} \quad \text{且} \quad (10)$$
$$\|x^* - x\|_{\ell_1} \leq C_0 \cdot \|x - x_S\|_{\ell_1}$$

這裡 x_S 表示去掉所有最大 S 元素後其餘置 0 的向量 x 。（如上所述，該成果應歸功於文獻 [17] 的第一作者，至今未公開發表，亦可參見文獻 [16] 和 [18]）

定理 2 的結論比定理 1 更爲有力。假如 x 是 S -稀疏，那麼 $x = x_S$ ，因此，可以準確恢復。但此新定理涉及到所有訊號。假如 x 並非 S -稀疏，那麼公式 (10) 主張恢復的訊號品質，與提前知道 x 中的 S 個最大值位置而決定直接量測的結果一樣好。換句話說，重構幾乎與提供了具有關於 x 的完整和完善知識，從爲我們提取 S 個最重要資訊片段的預言一樣好。

與我們早些的成果間的另一個明顯區別是其已確定，不涉及機率。假如我們有一個遵循定理假設的矩陣 A ，利用其即可以保證精確恢復所有 S -稀疏向量，對其他情況，實質上是恢復所有向量的 S 個最大元素，也就是說不可能失敗。

最不同尋常的是這些採樣協定允許感測器非常有效地捕捉一個稀疏訊號中的資訊，但並沒有試圖理解該訊號

the theorem, we may apply it, and we are then guaranteed to recover *all* sparse S -vectors exactly, and essentially the S -largest entries of *all* vectors otherwise; i.e., there is no probability of failure.

What is missing at this point is the relationship between S (the number of components one can effectively recover) obeying the hypothesis and m the number of measurements or rows of the matrix. To derive powerful results, we would like to find matrices obeying the RIP with values of S close to m . Can one design such matrices? In the next section, we will show that this is possible, but first we examine the robustness of CS vis a vis data corruption.

ROBUST SIGNAL RECOVERY FROM NOISY DATA

We are given noisy data as in (7) and use ℓ_1 minimization with relaxed constraints for reconstruction:

$$\min \|\tilde{x}\|_{\ell_1} \quad \text{subject to} \quad \|A\tilde{x} - y\|_{\ell_2} \leq \epsilon, \quad (11)$$

where ϵ bounds the amount of noise in the data. (One could also consider recovery programs such as the Dantzig selector [19] or a combinatorial optimization program proposed by Haupt and Nowak [20]; both algorithms have provable results in the case where the noise is Gaussian with bounded variance.) Problem (11) is often called the LASSO after [21]; see also [22]. To the best of our knowledge, it was first proposed in [8]. This is again a convex problem (a second-order cone program) and can be solved efficiently.

THEOREM 3 [16]

Assume that $\delta_{2S} < \sqrt{2} - 1$. Then the solution x^* to (11) obeys

$$\|x^* - x\|_{\ell_2} \leq C_0 \cdot \|x - x_S\|_{\ell_1} / \sqrt{S} + C_1 \cdot \epsilon \quad (12)$$

for some constants C_0 and C_1 . (Again, this theorem is unpublished as stated and is a variation on the result found in [16].)

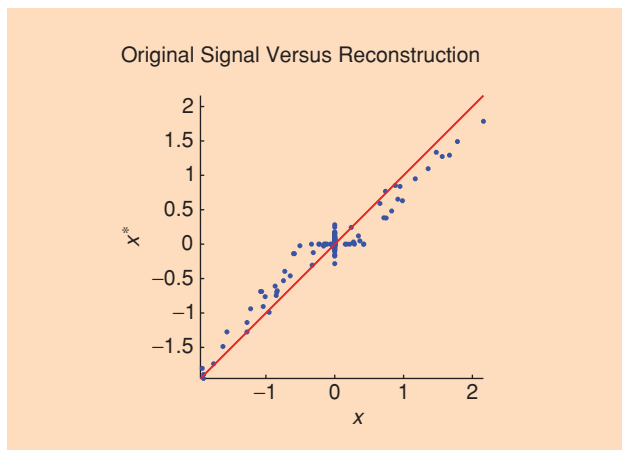
This can hardly be simpler. The reconstruction error is bounded by the sum of two terms. The first is the error which would occur if one had noiseless data. The second is just proportional to the noise level. Further, the constants C_0 and C_1 are typically small. With $\delta_{2S} = 1/4$ for example, $C_0 \leq 5.5$ and $C_1 \leq 6$. Figure 3 shows a reconstruction from noisy data.

This last result establishes CS as a practical and robust sensing mechanism. It works with all kinds of not necessarily sparse signals, and it handles noise gracefully. What remains to be done is to design efficient sensing matrices obeying the RIP. This is the subject of the next section.

RANDOM SENSING

Returning to the RIP, we would like to find sensing matrices with the property that column vectors taken from arbitrary subsets are nearly orthogonal. The larger these subsets, the better.

This is where randomness re-enters the picture. Consider the following sensing matrices: i) form A by sampling n col-



[FIG3] A signal x (horizontal axis) and its reconstruction x^* (vertical axis) obtained via (11). In this example, $n = 512$ and $m = 256$. The signal is 64-sparse. In the model (7), the sensing matrix has i.i.d. $N(0, 1/m)$ entries and z is a Gaussian white noise vector adjusted so that $\|Ax\|_{\ell_2} / \|z\|_{\ell_2} = 5$. Here, $\|x^* - x\|_{\ell_2} \approx 1.3 \cdot \epsilon$.

umn vectors uniformly at random on the unit sphere of \mathbb{R}^m ; ii) form A by sampling i.i.d. entries from the normal distribution with mean 0 and variance $1/m$; iii) form A by sampling a random projection P as in “Incoherent Sampling” and normalize: $A = \sqrt{n/m} P$; and iv) form A by sampling i.i.d. entries from a symmetric Bernoulli distribution ($P(A_{i,j} = \pm 1/\sqrt{m}) = 1/2$) or other sub-gaussian distribution. With overwhelming probability, all these matrices obey the RIP (i.e. the condition of our theorem) provided that

$$m \geq C \cdot S \log(n/S), \quad (13)$$

where C is some constant depending on each instance. The claims for i)–iii) use fairly standard results in probability theory; arguments for iv) are more subtle; see [23] and the work of Pajor and his coworkers, e.g., [24]. In all these examples, the probability of sampling a matrix not obeying the RIP when (13) holds is exponentially small in m . Interestingly, there are no measurement matrices and no reconstruction algorithm whatsoever which can give the conclusions of Theorem 2 with substantially fewer samples than the left-hand side of (13) [2], [3]. In that sense, using randomized matrices together with ℓ_1 minimization is a near-optimal sensing strategy.

One can also establish the RIP for pairs of orthobases as in “Incoherence and the Sensing of Sparse Signals.” With $A = R\Phi\Psi$ where R extracts m coordinates uniformly at random, it is sufficient to have

$$m \geq C \cdot S (\log n)^4, \quad (14)$$

for the property to hold with large probability; see [25] and [2]. If one wants a probability of failure no larger than $O(n^{-\beta})$ for some $\beta > 0$, then the best known exponent in (14) is five instead of four (it is believed that (14) holds with just $\log n$).

此時缺少的是遵循假設之 S (能有效恢復之分量數) 和矩陣的列向量或量測數 m 間的關係。為了推導出更有效的結果, 我們希望尋找滿足 RIP, 且 S 值接近 m 的矩陣。可以設計此矩陣嗎? 在下一節中, 我們將證明這是可能的, 但是首先我們來分析 CS 面對資料干擾時的穩健性。

雜訊資料的穩健訊號恢復

我們給定如公式 (7) 所示之雜訊資料, 採用不嚴格約束的極小化來重構:

$$\min \|\tilde{x}\|_{\ell_1} \quad \text{subject to} \quad \|A\tilde{x} - y\|_{\ell_2} \leq \varepsilon \quad (11)$$

這裡 ε 指明資料中雜訊值的界線 (這裡我們可以考慮恢復程式, 例如 Dantzig 選擇器 [19], 或 Haupt 和 Nowak 提出的組合最佳化程式 [20]; 兩個演算法在變異數有界的高斯雜訊情況下都有可證明的研究成果)。此問題 (11) 在文獻 [21] 發表後常被稱為 LASSO; 見文獻 [22]。就我們所知, 這最早在文獻 [8] 中提出。又是一個凸規劃問題 (二階錐規劃), 可以有效求解。

定理3 [16]

假設 $\delta_{2S} < \sqrt{2} - 1$, 則公式(11)的解 x^* 對於某一常數和服從

$$\|x^* - x\|_{\ell_2} \leq C_0 \cdot \|x - x_S\|_{\ell_1} / \sqrt{S} + C_1 \cdot \varepsilon \quad (12)$$

(此外, 如上所述的理論並未公開發表, 文獻 [16] 提出結果之變體)。

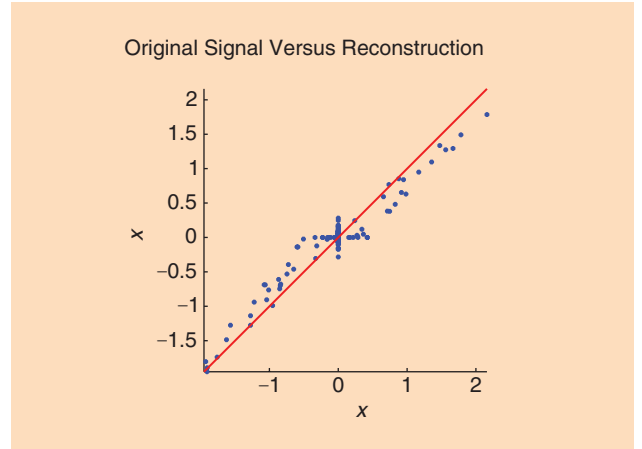
這幾乎不能再進一步簡化。重構誤差以兩項之和為界。第一個是無雜訊資料產生的誤差。第二個與雜訊水準成正比。進一步說, 常數 C_0 和 C_1 通常較小。例如 $\delta_{2S} = 1/4$, $C_0 \leq 5.5$ 且 $C_1 \leq 6$ 。圖 3 提出雜訊資料之重構。

這個最終結果證實 CS 是一個實用的和穩健的感測機制。幾乎對所有不一定稀疏之訊號都有效, 處理雜訊也很好。剩下要做的是來設計滿足 RIP 之有效感測矩陣。這是下一節的目標。

隨機感測

回到 RIP 的討論上來, 我們希望找到具有來自任意子集的行向量都幾乎正交這個性質之感測矩陣。這些子集規模越大越好。

因而隨機性再次參與其中。考慮如下感測矩陣 i) 均勻地在單位球面 \mathbb{R}^m 上隨機採樣 n 個行向量構成 A ; ii) 透過



【圖 3】訊號 (水平軸) 和透過公式 (11) 獲得的重構 (垂直坐標軸)。本例中, $n=512$, $m=256$ 。訊號是 64-稀疏的。在式 (7) 之模型中, 感測矩陣具有獨立同分佈於 $N(0, 1/m)$ 的項, z 是校準的高斯白色雜訊向量, 因此 $\|Ax\|_{\ell_2} / \|z\|_{\ell_2} = 5$ 。這裡 $\|x^* - x\|_{\ell_2} \approx 1.3 \cdot \varepsilon$ 。

對均值為 0, 變異數為 $1/m$ 之正態分佈進行獨立同分佈採樣構造 A ; iii) 如「非相干採樣」所述, 對隨機投影 P 採樣並將其歸一化: $A = \sqrt{n/m} P$ 來構造 A ; iv) 從對稱的伯努利 (Bernoulli) 分佈 ($P(A_{i,j} = \pm 1/\sqrt{m}) = 1/2$) 中或其他次高斯分佈上進行獨立同分佈採樣來構造 A 。如果

$$m \geq C \cdot S \log(n/S) \quad (13)$$

那麼所有這些矩陣都將以很大機率符合 RIP, 這裡 C 是取決於具體情況的某一常數。對於 i)-iii) 要求採用機率論中相當標準的結果; 關於 iv) 的參數則更為複雜些; 見文獻 [23] 和 Pajor 及其合作者的研究, 例如文獻 [24]。在所有這些範例中, 當滿足公式 (13) 時, 將採樣到一個不具備 RIP 矩陣的機率以指數形式減小到 m 。有趣的是, 沒有任何量測矩陣和重構演算法能用比(13)式左端更少的樣本提出定理 2 之結果[2][3]。在此意義上, 使用隨機化矩陣和 ℓ_1 極小化是近似最優的感測策略。

我們可以建立關於正交基對之 RIP, 正如在「非相干性和稀疏訊號的感測」所述。利用 $A = R\Phi\Psi^T$, 這裡 R 均勻隨機地抽取 m 個座標, 足以使:

$$m \geq C \cdot S(\log n)^4 \quad (14)$$

這是以較大機率具有該性質之充分條件; 見文獻 [25] 和 [2]。如果我們希望對於某些 $\beta > 0$, 失敗機率不超過 $O(n^{-\beta})$, 到目前為止, 式 (14) 中的已知最佳指數是 5 而非 4 (人們相信式 [14] 恰對 $\log n$ 成立)。這證明了能

This proves that one can stably and accurately reconstruct nearly sparse signals from dramatically undersampled data in an incoherent domain.

Finally, the RIP can also hold for sensing matrices $A = \Phi\Psi$, where Ψ is an arbitrary orthobasis and Φ is an $m \times n$ measurement matrix drawn randomly from a suitable distribution. If one fixes Ψ and populates Φ as in i)–iv), then with overwhelming probability, the matrix $A = \Phi\Psi$ obeys the RIP provided that (13) is satisfied, where again C is some constant depending on each instance. These random measurement matrices Φ are in a sense *universal* [23]; the sparsity basis need not even be known when designing the measurement system!

MATHEMATICAL AND COMPUTATIONAL METHODS COULD HAVE AN ENORMOUS IMPACT IN AREAS WHERE CONVENTIONAL HARDWARE DESIGN HAS SIGNIFICANT LIMITATIONS.

support is practically impossible. Whereas purely algebraic methods ignore the conditioning of information operators, having well-conditioned matrices, which are crucial for accurate estimation, is a central concern in CS as evidenced by the role played by the RIP.

WHAT IS COMPRESSIVE SAMPLING?

Data acquisition typically works as follows: massive amounts of data are collected only to be—in large part—discarded at the compression stage to facilitate storage and transmission. In the language of this article, one acquires a high-resolution pixel array f , computes the complete set of transform coefficients, encode the largest coefficients and discard all the others, essentially ending up with f_S . This process of massive data acquisition followed by compression is extremely wasteful (one can think about a digital camera which has millions of imaging sensors, the pixels, but eventually encodes the picture in just a few hundred kilobytes).

CS operates very differently, and performs as “if it were possible to directly acquire just the important information about the object of interest.” By taking about $O(S \log(n/S))$ random projections as in “Random Sensing,” one has enough information to reconstruct the signal with accuracy at least as good as that provided by f_S , the best S -term approximation—the best compressed representation—of the object. In other words, CS measurement protocols essentially translate analog data into an already compressed digital form so that one can—at least in principle—obtain super-resolved signals from just a few sensors. All that is needed after the acquisition step is to “decompress” the measured data.

There are some superficial similarities between CS and ideas in coding theory and more precisely with the theory and practice of Reed-Solomon (RS) codes [26]. In a nutshell and in the context of this article, it is well known that one can adapt ideas from coding theory to establish the following: one can uniquely reconstruct any S -sparse signal from the data of its first $2S$ Fourier coefficients, $y_k = \sum_{t=0}^{n-1} x_t e^{-i 2\pi kt/n}$, $k = 0, 1, 2, \dots, 2S - 1$, or from any set of $2S$ consecutive frequencies for that matter (the computational cost for the recovery is essentially that of solving an $S \times S$ Toeplitz system and of taking an n -point fast Fourier transform). Does this mean that one can use this technique to sense compressible signals? The answer is negative and there are two main reasons for this. First, the problem is that RS decoding is an algebraic technique, which cannot deal with nonsparse signals (the decoding finds the support by rooting a polynomial);

second, the problem of finding the support of a signal—even when the signal is exactly sparse—from its first $2S$ Fourier coefficients is extraordinarily ill posed (the problem is the same as that of extrapolating a high degree polynomial from a small number of highly clustered values). Tiny perturbations of these coefficients will give completely different answers so that with finite precision data, reliable estimation of the sup-

APPLICATIONS

The fact that a compressible signal can be captured efficiently using a number of *incoherent* measurements that is proportional to its information level $S \ll n$ has implications that are far reaching and concern a number of possible applications:

- Data compression. In some situations, the sparse basis Ψ may be unknown at the encoder or impractical to implement for data compression. As we discussed in “Random Sensing,” however, a randomly designed Φ can be considered a universal encoding strategy, as it need not be designed with regards to the structure of Ψ . (The knowledge and ability to implement Ψ are required only for the decoding or recovery of f .) This universality may be particularly helpful for distributed source coding in multi-signal settings such as sensor networks [27]. We refer the reader to articles by Haupt et al. and Goyal et al. elsewhere in this issue for related discussions.
- Channel coding. As explained in [15], CS principles (sparsity, randomness, and convex optimization) can be turned around and applied to design fast error correcting codes over the reals to protect from errors during transmission.
- Inverse problems. In still other situations, the only way to acquire f may be to use a measurement system Φ of a certain modality. However, assuming a sparse basis Ψ exists for f that is also incoherent with Φ , then efficient sensing will be possible. One such application involves MR angiography [1] and other types of MR setups [28], where Φ records a subset of the Fourier transform, and the desired image f is sparse in the time or wavelet domains. Elsewhere in this issue, Lustig et al. discuss this application in more depth.
- Data acquisition. Finally, in some important situations the full collection of n discrete-time samples of an analog signal may be difficult to obtain (and possibly difficult to subsequently compress). Here, it could be helpful to design physical sampling devices that directly record discrete, low-rate incoherent measurements of the incident analog signal.

The last of these applications suggests that mathematical and computational methods could have an enormous impact

夠從非相干域明顯欠採樣資料中近乎穩定和精確地重構稀疏訊號。

最後，感測矩陣 $A = \Phi\Psi$ 也可以具有 RIP，這裡 Ψ 是任一個正交基， Φ 是一個從適當分佈中隨機抽取的 $m \times n$ 大小的感測矩陣。假設 Ψ 固定，如 i)-iv) 所述構造 Φ ，如果滿足公式 (13)，那麼矩陣 $A = \Phi\Psi$ 以很大的機率服從 RIP，這裡， C 是取決於具體情況的某一常數。在某種意義上，這些隨機感測矩陣 Φ 是普適的 [23]；當設計量測系統時，甚至不需要事先知道稀疏基！

什麼是壓縮取樣？

通常資料擷取按如下方式實施：先收集大量資料，其中大部分在之後的壓縮階段被丟棄，以利於儲存和傳輸。在本文表述中，採集高解析度像素陣列 f ，計算全部轉換係數集合並對最大係數進行編碼，丟棄其它所有係數，本質上以 f_s 結束。大量資料在採集過程後緊接著做壓縮是極端的浪費（我們可以想像一個影像感測器高達百萬像素的數碼相機，最終編碼為數百 KB 的照片）。

CS 運算則完全不同，其表現為「如果可能直接擷取的話，僅擷取於感興趣目標相關的重要資訊」。藉由約 $O(S \log(n/S))$ 個隨機投影，正如「隨機感測」中一樣，資訊足以精準重構訊號，至少如 f_s 提供的目標之最佳 S -項逼近——最佳壓縮表示一樣好。換句話說，CS 量測協定本質上將類比資料轉化為到壓縮的數位形式，如此這般，可以——至少在原則上——僅從少數感測器上得到超分辨訊號。擷取後所有的必要步驟是「壓縮」量測資料。

CS 與編碼理論的思路些微類似，準確地說是里德所羅門 Reed-Solomon (RS) 碼 [26] 的理論和應用。就本文的內容簡而言之，眾所周知，我們可以採用編碼理論之思路，設立如下：可以從其最開始的 $2S$ 個傅立葉係數資料中， $y_k = \sum_{i=0}^{n-1} x_i e^{-i2\pi kt/n}$, $k = 0, 1, 2, \dots, 2S - 1$ ，或者從 $2S$ 個連續頻率之任何集合中唯一重構任何 S -稀疏訊號。（恢復的計算成本本質上是求解 $S \times S$ Toeplitz 系統和計算 n 點之快速傅立葉轉換）。這是否意味著使用該技術可以感測可壓縮的訊號？答案是否定的，基於如下兩個原因。首先，RS 編碼作為一種代數方法，不能處理非稀疏訊號（解碼是透過對多項式求根來尋找支撐集）；其次，從其前 S 個傅立葉係數中尋找訊號支撐集的問題——即使訊號恰是稀疏——是嚴格不適定（該問題與從少數高度集中的數值中推導而得的高次多項式形式相同）。這些係數

的微小擾動會導致完全不同的答案，所以藉由有限精度之資料，提供可靠的估計實際上是不可能的。鑒於純粹的代數方法有忽略資訊運算元之調節作用，所得到的良態矩陣，這是精確重構之關鍵，是 CS 關注的中心問題，正如由 RIP 扮演的角色所證明的一樣。

陣，這是精確重構之關鍵，是 CS 關注的中心問題，正如由 RIP 扮演的角色所證明的一樣。

數學方法和計算方法在傳統硬體設計受限的領域產生巨大的影響

應用

可壓縮訊號可以有效捕捉與其資訊層 $S \ll n$ 成比例的若干非相干量測值，這具有深遠的意義，並關係到一些可能的應用：

- 資料壓縮。在某些情況下，編碼器中的稀疏基 Ψ 可能不是已知的，或者對於資料壓縮的實現不切實際。正如我們在「隨機感測」中討論的一樣，隨機設計的 Φ 被認為是一個普適的編碼策略，因為不需要設計關於 Ψ 之結構（ Ψ 所具有的知識和能力僅需用在 f 的解碼和恢復中）。如此普適性對於多訊號裝置中的分散式源編碼特別有利，例如感測器網路 [27]。我們希望讀者參閱 Haupt 等人 and Goyal 等人在它處發表的對本問題之相關討論。

- 通道編碼。如文獻 [15] 中所述，CS 原理（稀疏性，隨機性和凸型最佳化）可以轉化和應用於設計快速錯誤修正碼，以避免傳輸中出現錯誤訊號。

- 反問題。還有其它情況，擷取 f 之唯一方法可能就是使用某種形式的量測值系統 Φ 。然而，假設存在一個關於 f 的稀疏基 Ψ ，而且與 Φ 非相干，那麼有可能發生有效感測。這類應用包含了 MR 血管造影術 [1] 和其它類型的 MR 設備 [28]，此處 Φ 記錄傅立葉轉換的子集，希望得到的圖像 f 在時域或者小波域是稀疏的。在本問題的其它領域，Lustig 等人對其有更深入的討論。

- 資料擷取。最後，在一些重要情況下，難以完全收集類比訊號的 n 個離散時間樣本（可能難以實施後續壓縮）。此時，設計實體採樣裝置，有助於直接記錄類比訊號之離散、低速率的非相干量測。

最後這些應用表明數學和計算的方法在這個領域產生巨大的影響，而在該領域中，傳統硬體設計存在重要局限。例如，傳統上採用 CCD 或 CMOS 技術之影像設備在本質上受限於可視光譜。然而，CS 相機利用數位微鏡面陣列採集非相干之量測值（僅需要一個光敏元件取代百萬像素），可極大地拓展其性能（關於此問題，參見文獻 [29] 和由 Duarte 等人撰寫的文章）。

in areas where conventional hardware design has significant limitations. For example, conventional imaging devices that use CCD or CMOS technology are limited essentially to the visible spectrum. However, a CS camera that collects incoherent measurements using a digital micromirror array (and requires just one photosensitive element instead of millions) could significantly expand these capabilities. (See [29] and an article by Duarte et al. in this issue.)

Along these same lines, part of our research has focused on advancing devices for “analog-to-information” (A/I) conversion of high-bandwidth signals (see also the article by Healy et al. in this issue). Our goal is to help alleviate the pressure on conventional ADC technology, which is currently limited to sample rates on the order of 1 GHz. As an alternative, we have proposed two specific architectures for A/I in which a discrete, low-rate sequence of incoherent measurements can be acquired from a high-bandwidth analog signal. To a high degree of approximation, each measurement y_k can be interpreted as the inner product $\langle f, \varphi_k \rangle$ of the incident analog signal f against an analog measurement waveform φ_k . As in the discrete CS framework, our preliminary results suggest that analog signals obeying a sparse or compressible model (in some analog dictionary Ψ) can be captured efficiently using these devices at a rate proportional to their information level instead of their Nyquist rate. Of course, there are challenges one must address when applying the discrete CS methodology to the recovery of sparse analog signals. A thorough treatment of these issues would be beyond the scope of this short article and as a first cut, one might simply accept the idea that in many cases, discretizing/sampling the sparse dictionary allows for suitable recovery. Our two architectures are as follows:

1) **Nonuniform Sampler (NUS)**. Our first architecture simply digitizes the signal at randomly or pseudo-randomly sampled time points. That is, $y_k = f(t_k) = \langle f, \delta_{t_k} \rangle$. In effect, these

time points are obtained by jittering nominal (low-rate) sample points located on a regular lattice. Due to the incoherence between spikes and sines, this architecture can be used to sample signals having sparse frequency spectra far below their Nyquist rate. There are of course tremendous benefits associated with a reduced sampling rate, as this provides added circuit settling time and has the effect of reducing the noise level.

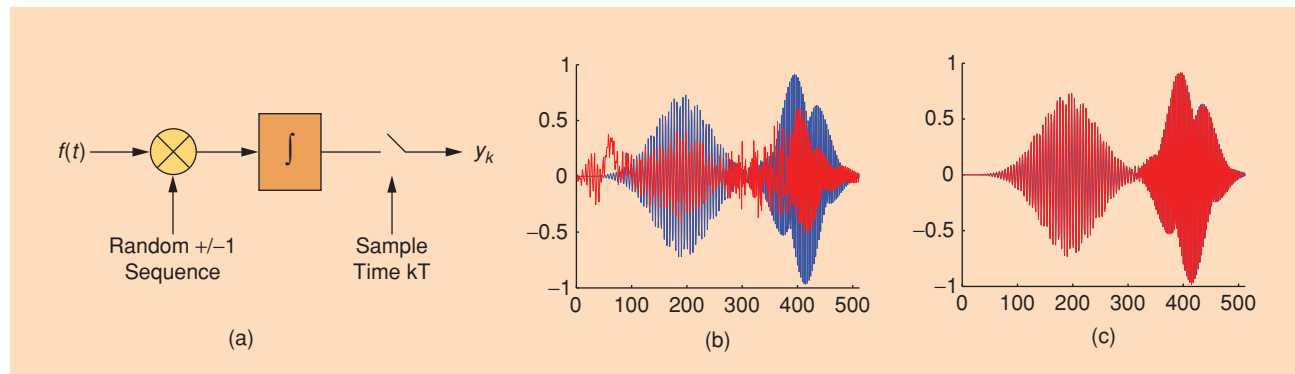
2) **Random Modulation Preintegration (RMPI)**. Our second architecture is applicable to a wider variety of sparsity domains, most notably those signals having a sparse signature in the time-frequency plane. Whereas it

may not be possible to digitize an analog signal at a very high rate rate, it may be quite possible to change its polarity at a high rate. The idea of the RMPI architecture [see Figure 4(a)] is then to multiply the signal by a pseudo-random sequence of ± 1 s, integrate the product over time windows, and digitize the integral at the end of each time interval. This is a parallel architecture and one has several of these random multiplier-integrator pairs running in parallel using distinct sign sequences. In effect, the RMPI architecture correlates the signal with a bank of sequences of ± 1 , one of the random CS measurement processes known to be universal, and therefore the RMPI measurements will be incoherent with any fixed time-frequency dictionary such as the Gabor dictionary described below.

For each of the above architectures, we have confirmed numerically (and in some cases physically) that the system is robust to circuit nonidealities such as thermal noise, clock timing errors, interference, and amplifier nonlinearities.

The application of A/I architectures to realistic acquisition scenarios will require continued development of CS algorithms and theory. To highlight some promising recent directions, we conclude with a final discrete example. We take f to be a one-dimensional signal of length $n = 512$ that contains two modulated pulses [see the blue curve in Figure 4(b)] From this signal, we

PART OF OUR RESEARCH HAS FOCUSED ON ADVANCING DEVICES FOR “ANALOG-TO-INFORMATION” (A/I) CONVERSION OF HIGH-BANDWIDTH SIGNALS.



[FIG4] Analog-to-information conversion. (a) Random modulation preintegration (RMPI) system. (b) Original two-pulse signal (blue) and reconstruction (red) via ℓ_1 synthesis from random ± 1 measurements. (c) Two-pulse signal and reconstruction via reweighted ℓ_1 analysis.

在這方面，我們的部分研究關注於推進高頻寬訊號的「類比到資訊」(A/I) 設備之發展 (參見Healy 等就此問題撰寫的文章)。我們的目標是緩和傳統 ADC 技術之壓力，當前其受到約為 1GHz 採樣速率的限制。作為替代方案，我們提出關於 A/I 的兩個特定架構，可以從高頻寬類比訊號中獲取離散、低速率的非相干之量測序列。為了得到高度逼近，每個量測值 y_k 被解釋為注入類比訊號 f 與一個類比量測波形 f_k 之內積 $\langle f, f_k \rangle$ 。正如在離散 CS 框架中一樣，初步結果表明滿足稀疏或可壓縮模型 (在某個類比字典上) 之類比訊號可以透過這些設備，以與其資訊層成比例的速率，並非奈奎斯特速率來有效擷取。當然，利用離散 CS 方法論恢復稀疏類比訊號時還存在有待克服的挑戰。有關這些問題的全面論述將超過本短文的範圍，我們可以簡單地接受其思想，即在許多情況下，離散化/採樣稀疏字典將允許適當恢復。我們的兩個架構如下所示：

**我們的部分研究關注於推進
高頻寬訊號的「類比到資訊」
(A/I) 設備之發展**

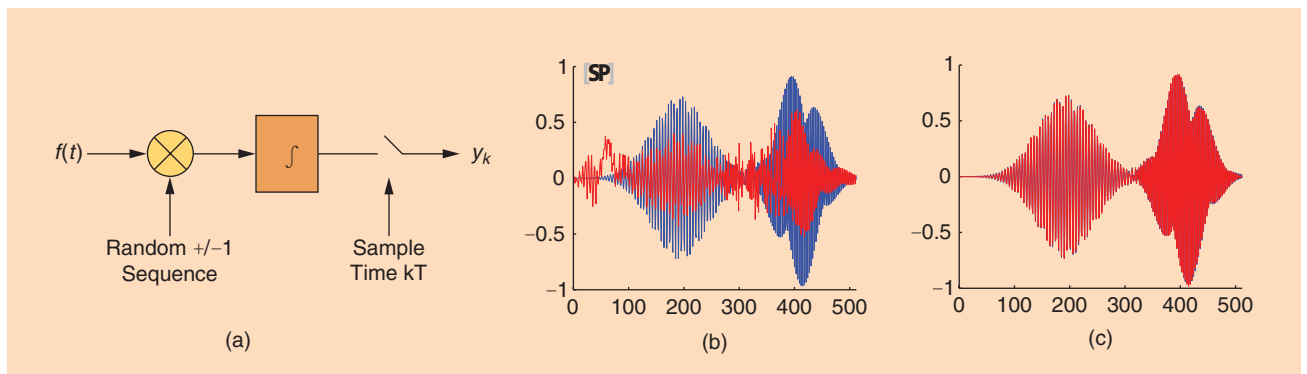
1) **非均勻採樣器 (NUS)**。我們的第一個構架只不過是在隨機或偽隨機採樣時間點上的訊號數位化，即 $y_k = f(t_k) = \langle f, d_k \rangle$ 。實際上這些隨機或者偽隨機時間點是透過抖動位於規則格子上的 (低採樣率) 採樣點擷取。由於尖峰與正弦訊號間的非相干性，因此此架構可用於採樣頻譜稀疏且遠低於奈奎斯特頻率的訊號。降低採樣率當然帶來的好處極多，因為這提供更多的電路穩定時間，並具有減小雜訊水準的作用。

2) **隨機調變預積分 (RMPI)**。我們的第二種構架可用於更多種的稀疏域，最明顯的範例就是那些在時間頻率平面上有稀疏特徵之訊號。然而，我們不可能以高採樣速率將類比訊號數位化，但極有可能改變其在高採樣速率

下的極性。RMPI 架構的思想 (見圖 4(a)) 即是以 $\pm 1s$ 的偽隨機序列乘以訊號，在時窗上對乘積進行積分，數位化時段末之積分值。這是一個並行架構，有若干這樣的乘法器積分器對，使用完全不同的符號序列進行並行運算。實際上，RMPI 架構將訊號同 ± 1 的符號序列庫聯繫起來，此為通用 CS 感測過程方法之一，因此 RMPI 量測與任何確定的時頻字典都非相干，例如下述之 Gabor 字典。

對於上述每個構架，我們都用數值法進行論證 (有些狀況還有用物理方法)，系統對於電路的非理想狀態，如熱雜訊、時鐘誤差、干擾及放大器非線性等情況表現穩健。

將 A/I 構架應用於實際採集方案時還需要 CS 演算法和理論之進一步發展。為了強調一些未來的發展趨勢，我們最後以一個離散的範例結束。我們擷取長度為 $n = 512$ 的一維訊號 f ，包含兩個調製脈衝 (見圖 4(b) 中的藍色曲線)，從此訊號中，我們啟用獨立同分佈的伯努利 ± 1 輸入組成感測矩陣 Φ ，採集 $m = 30$ 次量測值。相應於欠採樣因數大於 17，少量資料並不合理。為了重構訊號，我們慮及 Gabor 字典 Ψ ，由 Gaussian 窗限時，且具有不同位置和純量的各種正弦波構成。總的來說，字典是近似 $43 \times$ 過完備，不含由 f 組成的兩個脈衝。圖 4(b) 中的紅色曲線提出了使 $\|x\|_{\ell_1}$ 極小化的結果，以致 $y = \Phi \Psi x$ 。重構顯示出顯著的人為效應，得到 $\|f - \hat{f}\|_{\ell_2} / \|f\|_{\ell_2} \approx 0.67$ 。然而實際上我們透過對 ℓ_1 範數的恢復程式做兩個更改以消除這些人為效應。首先，改為將 $\|\Psi^* \tilde{f}\|_{\ell_1}$ 最小化 (當 Ψ 是一個正交基時這一改變不起作用)。其次，在得到 \hat{f} 的估計後，我們將 ℓ_1 范數重新加權，再重複重構過程，對那些預計較大的係數採用較低的懲罰值。圖 4(c) 提出了四次



[圖 4] 類比資訊轉換。(a) 隨機調製預積分(RMPI)系統。(b) 原始的雙尖峰訊號 (藍色) 和由隨機 ± 1 量測經 ℓ_1 合成的重構結果 (紅色)。(c) 雙尖峰訊號和由重新加權的 ℓ_1 分析所得到的重構結果。

collect $m = 30$ measurements using an $m \times n$ measurement matrix Φ populated with i.i.d. Bernoulli ± 1 entries. This is an unreasonably small amount of data corresponding to an under-sampling factor of over 17. For reconstruction we consider a Gabor dictionary Ψ that consists of a variety of sine waves time limited by Gaussian windows, with different locations and scales. Overall the dictionary is approximately $43\times$ overcomplete and does not contain the two pulses that comprise f . The red curve in Figure 4(b) shows the result of minimizing $\|x\|_{\ell_1}$ such that $y = \Phi\Psi x$. The reconstruction shows pronounced artifacts, and we see $\|f - f^*\|_{\ell_2}/\|f\|_{\ell_2} \approx 0.67$. However, we can virtually eliminate these artifacts by making two changes to the ℓ_1 recovery program. First, we instead minimize $\|\Psi^* \tilde{f}\|_{\ell_1}$ subject to $y = \Phi \tilde{f}$. (This variation has no effect when Ψ is an orthobasis.) Second, after obtaining an estimate f^* , we reweight the ℓ_1 norm and repeat the reconstruction, with a lower penalty applied to those coefficients we anticipate to be large. Figure 4(c) shows the result after four iterations of reweighting; we see $\|f - f^*\|_{\ell_2}/\|f\|_{\ell_2} \approx 0.022$. We refer the reader to [30] for more information on these directions. The point here is that even though the amount of data is ridiculously small, one has nevertheless captured most of the information contained in the signal. This, in a nutshell, is why CS holds such great promise.

AUTHORS

Emmanuel J. Candès (emmanuel@acm.caltech.edu) received his B. Sc. degree from the École Polytechnique, France, in 1993 and the Ph.D. degree in statistics from Stanford University in 1998. He is the Ronald and Maxine Linde Professor of Applied and Computational Mathematics at the California Institute of Technology. His research interests are in computational harmonic analysis, statistical estimation and detection, signal processing, scientific computing, inverse problems, and mathematical optimization. He received the Third Popov Prize in Approximation Theory in 2001, and the DOE Young Investigator Award in 2002. He was an Alfred P. Sloan Research Fellow in 2001. He has given plenary addresses at major international conferences, including ICIAM 2007 and ICIP 2007. In 2005, he was awarded the James H. Wilkinson Prize in Numerical Analysis and Scientific Computing by SIAM. He received the NSF 2006 Alan T. Waterman Medal.

Michael B. Wakin (wakin@umich.edu) received the B.S. degree in electrical engineering and the B.A. degree in mathematics in 2000 (summa cum laude), the M.S. degree in electrical engineering in 2002, and the Ph.D. degree in electrical engineering in 2007, all from Rice University. From 2006–2007, he was an NSF Mathematical Sciences postdoctoral research fellow in the Department of Applied and Computational Mathematics at the California Institute of Technology, and he is currently an assistant professor in the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor. His research interests include sparse, geometric, and manifold-based models for signal and image processing, approximation, compression, compressive sampling, and dimensionality reduction.

REFERENCES

- [1] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [2] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [3] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [4] D.S. Taubman and M.W. Marcellin, *JPEG 2000: Image Compression Fundamentals, Standards and Practice*. Norwell, MA: Kluwer, 2001.
- [5] D.L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.
- [6] R. Coifman, F. Geshwind, and Y. Meyer, "Noiselets," *Appl. Comput. Harmon. Anal.*, vol. 10, no. 1, pp. 27–44, 2001.
- [7] J.F. Claerbout and F. Muir, "Robust modeling with erratic data," *Geophys. Mag.*, vol. 38, no. 5, pp. 826–844, Oct. 1973.
- [8] F. Santosa and W.W. Symes, "Linear inversion of band-limited reflection seismograms," *SIAM J. Sci. Statist. Comput.*, vol. 7, no. 4, pp. 1307–1330, 1986.
- [9] J. Tropp and A.C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [10] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Prob.*, vol. 23, no. 3, pp. 969–985, 2007.
- [11] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, Atlanta, GA, vol. 2, 1996, pp. 1689–1692.
- [12] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002.
- [13] A. Gilbert, S. Muthukrishnan, and M. Strauss, "Improved time bounds for near-optimal sparse Fourier representation," in *Proc. Wavelets XI SPIE Optics Photonics*, San Diego, CA, 2005.
- [14] D.L. Donoho and P.B. Stark, "Uncertainty principles and signal recovery," *SIAM J. Appl. Math.*, vol. 49, no. 3, pp. 906–931, 1989.
- [15] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [16] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
- [17] E.J. Candès, "Lectures on compressive sampling and frontiers in signal processing," *The Institute for Mathematics and its Applications*. University of Minnesota, June 2007 [Online]. Available: <http://www.ima.umn.edu/2006-2007/ND6.4-15.07/abstracts.html>
- [18] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best k -term approximation," 2006, Preprint.
- [19] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n ," submitted for publication.
- [20] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 4036–4048, 2006.
- [21] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.
- [22] S.S. Chen, D.L. Donoho, and M.A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.
- [23] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," submitted for publication.
- [24] S. Mendelson, A. Pajor, and N. Tomczak-Jaegermann, "Uniform uncertainty principle for Bernoulli and sub-gaussian ensembles," 2006, Preprint.
- [25] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," submitted for publication.
- [26] R.E. Blahut, *Algebraic Codes for Data Transmission*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [27] D. Baron, M.B. Wakin, M.F. Duarte, S. Sarvotham, and R.G. Baraniuk, "Distributed compressed sensing," 2005, Preprint.
- [28] M. Lustig, D.L. Donoho, and J.M. Pauly, "Rapid MR imaging with compressed sensing and randomly under-sampled 3DFT trajectories," in *Proc. 14th Ann. Meeting ISMRM*, Seattle, WA, May 2006.
- [29] D. Takhar, V. Bansal, M. Wakin, M. Duarte, D. Baron, K.F. Kelly, and R.G. Baraniuk, "A compressed sensing camera: New theory and an implementation using digital micromirrors," in *Proc. Comp. Imaging IV SPIE Electronic Imaging*, San Jose, CA, 2006.
- [30] E.J. Candès, M.B. Wakin, and S.P. Boyd, "Enhancing sparsity by reweighting ℓ_1 ," Tech. Rep., California Institute of Technol., 2007 [Online]. Available: <http://www.acm.caltech.edu/~emmanuel/publications.html>

重新加權反覆運算後的結果，得到 $\|f - \hat{f}\|_2 / \|f\|_2 \approx 0.022$ 。關於這些方面的更多資訊，請讀者參閱文獻 [30]。這裡的重點是，即使資料量小到令人吃驚的程度，我們仍可捕捉到訊號所包含的大部分資訊。簡而言之，這就是 CS 極有前途的緣故。

作者簡介

Emmanuel J. Candès (emmanuel@acm.caltech.edu) received his B. Sc. degree from the École Polytechnique, France, in 1993 and the Ph.D. degree in statistics from Stanford University in 1998. He is the Ronald and Maxine Linde Professor of Applied and Computational Mathematics at the California Institute of Technology. His research interests are in computational harmonic analysis, statistical estimation and detection, signal processing, scientific computing, inverse problems, and mathematical optimization. He received the Third Popov Prize in Approximation Theory in 2001, and the DOE Young Investigator Award in 2002. He was an Alfred P. Sloan Research Fellow in 2001. He has given plenary addresses at major international conferences, including ICIAM 2007 and ICIP 2007. In 2005, he was awarded the James H. Wilkinson Prize in Numerical Analysis and Scientific Computing by SIAM. He received the NSF 2006 Alan T. Waterman Medal.

Michael B. Wakin (wakin@umich.edu) received the B.S. degree in electrical engineering and the B.A. degree in mathematics in 2000 (summa cum laude), the M.S. degree in electrical engineering in 2002, and the Ph.D. degree in electrical engineering in 2007, all from Rice University. From 2006–2007, he was an NSF Mathematical Sciences postdoctoral research fellow in the Department of Applied and Computational Mathematics at the California Institute of Technology, and he is currently an assistant professor in the Department of Electrical Engineering and Computer Science at the University of Michigan, Ann Arbor. His research interests include sparse, geometric, and manifold-based models for signal and image processing, approximation, compression, compressive sampling, and dimensionality reduction.

參考文獻

- [1] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [2] E. Candès and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [3] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [4] D.S. Taubman and M.W. Marcellin, *JPEG 2000: Image Compression Fundamentals, Standards and Practice*. Norwell, MA: Kluwer, 2001.
- [5] D.L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.
- [6] R. Coifman, F. Geshwind, and Y. Meyer, "Noiselets," *Appl. Comput. Harmon. Anal.*, vol. 10, no. 1, pp. 27–44, 2001.
- [7] J.F. Claerbout and F. Muir, "Robust modeling with erratic data," *Geophys. Mag.*, vol. 38, no. 5, pp. 826–844, Oct. 1973.
- [8] F. Santosa and W.W. Symes, "Linear inversion of band-limited reflection seismograms," *SIAM J. Sci. Statist. Comput.*, vol. 7, no. 4, pp. 1307–1330, 1986.
- [9] J. Tropp and A.C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

- [10] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Prob.*, vol. 23, no. 3, pp. 969–985, 2007.
- [11] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proc. IEEE Int. Conf. Acoustics Speech Signal Processing*, Atlanta, GA, vol. 2, 1996, pp. 1689–1692.
- [12] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1417–1428, June 2002.
- [13] A. Gilbert, S. Muthukrishnan, and M. Strauss, "Improved time bounds for near-optimal sparse Fourier representation," in *Proc. Wavelets XI SPIE Optics Photonics*, San Diego, CA, 2005.
- [14] D.L. Donoho and P.B. Stark, "Uncertainty principles and signal recovery," *SIAM J. Appl. Math.*, vol. 49, no. 3, pp. 906–931, 1989.
- [15] E. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [16] E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.
- [17] E.J. Candès, "Lectures on compressive sampling and frontiers in signal processing," *The Institute for Mathematics and its Applications*. University of Minnesota, June 2007 [Online]. Available: <http://www.ima.umn.edu/2006-2007/ND6.4-15.07/abstracts.html>
- [18] A. Cohen, W. Dahmen, and R. DeVore, "Compressed sensing and best -term approximation," 2006, Preprint.
- [19] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when is much larger than ," submitted for publication.
- [20] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 4036–4048, 2006.
- [21] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Stat. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.
- [22] S.S. Chen, D.L. Donoho, and M.A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.
- [23] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," submitted for publication.
- [24] S. Mendelson, A. Pajor, and N. Tomczak-Jaegermann, "Uniform uncertainty principle for Bernoulli and sub-gaussian ensembles," 2006, Preprint.
- [25] M. Rudelson and R. Vershynin, "On sparse reconstruction from Fourier and Gaussian measurements," submitted for publication.
- [26] R.E. Blahut, *Algebraic Codes for Data Transmission*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [27] D. Baron, M.B. Wakin, M.F. Duarte, S. Sarvotham, and R.G. Baraniuk, "Distributed compressed sensing," 2005, Preprint.
- [28] M. Lustig, D.L. Donoho, and J.M. Pauly, "Rapid MR imaging with compressed sensing and randomly under-sampled 3DFT trajectories," in *Proc. 14th Ann. Meeting ISMRM*, Seattle, WA, May 2006.
- [29] D. Takhar, V. Bansal, M. Wakin, M. Duarte, D. Baron, K.F. Kelly, and R.G. Baraniuk, "A compressed sensing camera: New theory and an implementation using digital micromirrors," in *Proc. Comp. Imaging IV SPIE Electronic Imaging*, San Jose, CA, 2006.
- [30] E.J. Candès, M.B. Wakin, and S.P. Boyd, "Enhancing sparsity by reweighting ," Tech. Rep., California Institute of Technol., 2007 [Online]. Available: <http://www.acm.caltech.edu/emmanuel/publications.html>

元件鑑識技術

理論，方法及應用



視

覺感測器技術在近幾十年中取得了巨大的發展，數位設備變得無處不在。由各種影像裝置擷取的數位影像被大量應用，從軍事和偵察到醫學診斷和消費型攝影產品。因此，在成像技術的快速

發展和廣泛應用中，出現了一系列新鑑識問題。例如，不僅可以很容易地查出硬體和軟體元件是什麼類型，而且包括這些設備內部採用的參數嗎？某一幅特定數位圖像是用哪種成像感測器，或哪個品牌的感測器擷取的呢？該圖像是如何擷取的？使用的是數位相機、拍照手機、圖像掃描器，還是經過圖像編輯軟體合成的呢？擷取此圖像後是否經過了一些操弄？是真實的呢，還是經過了某種方式的篡改？其中是否包含隱藏資訊或者加密資料？許多鑑識問題

都與數位圖像的來源及其創建過程有關。得自於這種分析的證據可為司法、治安和情報機構提供有用的鑑識資訊。瞭解圖像採集技術，也有助於解答關於圖像擷取後，可能經過的附加處理之特質等更為深入的鑑識問題。

當前有許多處理這些問題的方法。在本文中，我們將研究主流技術之一：基於元件鑑識的技術來解答這些鑑識問題。元件鑑識之目標是識別擷取資料時不同設備元件中之演算法和參數。元件鑑識分析的工作原理是：尋找數位圖像在一系列資訊處理中，經由不同處理塊時所留下的固有指紋痕跡，並使用這些痕跡估計元件參數。透過固有指紋的識別建置元件參數分析，元件鑑識為許多鑑識問題之處理提供一個框架，例如發現設備技術侵權，保護智慧財產權和識別採集設備等等。

保護圖像設備的智慧財產權成為近年來首要關心的事情，電子成像行業的激烈競爭導致法庭上提交侵權案例數量的增加。勝訴方獲取的賠償金額也急速增加，有時高達數十億美金。眾所周知，專利是保護智慧財產權的一種強大工具。然而，隨著現代尖端工具的發展，影像裝置產品的專利侵權變得易如反掌，察覺困難，甚至難以在法庭上舉證。實現侵權分析的一般方法是對產品的設計和實作進行調查，透過若干類型的反向工程尋找該產品與現有專利主張之間的相似性。然而，該方法非常繁瑣和低效，在許多案例中可能包含牽涉調查設備軟體模組中低階組合語言的逐條比對。元件鑑識藉由在一系列資訊處理中每個元件上進行演算法和參數之識別，為侵權/授權鑑識提供一個系統性的方法論，由而保護智慧財產權。

元件鑑識還提供了用以確定圖像可信度和影像裝置的基礎。隨著操弄多媒體資料工具的快速發展，當圖像在新聞、偵察和司法應用上作為決定性的證據時，內容和採集設備的完整性就愈形重要。例如，關於設備中硬體/軟體模組及其參數等資訊有助於建立設備的識別系統。這些系統可為司法和情治機構提供關於採集圖像時所使用的是哪種設備或哪個商標/型號等有用的擷取鑑識資訊。另外，元件鑑識有助於確定一個實體模型，以確定直接從設備擷取之圖像特徵，相應地推動篡改鑑識，進而確定圖像離開設備後是否經過了任何的附加編輯和資料處理。

視可用的輸入之特質而定，元件鑑識主要可在三類情境下進行。在侵入式鑑識情境下，鑑識分析者具有取用設備的許可，並且可拆解該設備，析出每個元件，並提出計算個別元件參數的方法。在半非侵入式鑑識情境下，分析仍然具有取用設備的許可，但不許拆解該設備，那麼他/她可以設計適當的輸入資料輸入設備，以便收集關於處理技術和個別元件參數等鑑識證據。在完全的非侵入式鑑識情境下，鑑識分析者僅可基於設備的樣本輸出資料進行元件參數的估計。

本文中，我們將用視覺感測器和數碼相機擷取的圖像來論證元件鑑識，此時這些技術將進行適當修改和擴展為其它類型的擷取模型和感測技術。我們評估了各種相機

元件的參數估計方法。我們展示所計算出的參數可用來估計相機技術的相似性，用來在侵權/授權中提供能說明問題之線索、識別捕捉所討論圖像之照相機類型和品牌/型號，並構建一個真實背景模型來說明對檢測圖像內容的操弄。

數位影像裝置的系統模型

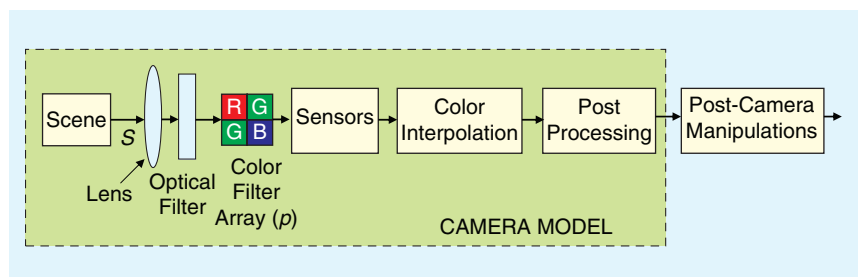
本章節中，我們評述一系列資訊處理中用以測試各種元件的數碼相機之圖像擷取模型。如圖 1 所示之圖像擷取模型，來自場景的光源傳過一個透鏡和光學濾鏡，最終由感測器陣列記錄下來。大多數相機採用彩色濾鏡陣列 (CFA) 捕捉來自真實世界場景的資訊。CFA 是感測器上的一個薄膜，選擇性地允許光源的某個特定成份通過薄膜，到達感測器 [1]。為了便於討論，定義 S 為被相機捕捉的真實世界場景， P 表示為 CFA 模式矩陣。 $S(x, y, c)$ 表示一個大小為 $H \times W \times C$ 的像素值三維 (3D) 陣列，這裡 H 和 W 分別表示圖像的高度和寬度， $C = 3$ 表示顏色元件的數目 (紅、綠和藍)。CFA 採樣使真實世界的場景轉化為一個 3D 矩陣 S_p 的形式

$$S_p(x, y, c) = \begin{cases} S(x, y, c) & \text{if } p(x, y) = c \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

在記錄下從 CFA 獲得的資料後，式 (1) 中對應於點 $S_p(x, y, c) = 0$ 的中間像素值透過一種眾所周知的處理運算——色彩插值或內插 [2]，利用其鄰域像素值進行插值並得到 S_p^0 。插值後，對應於紅、綠、藍元件的三個圖像透過一個後處理階段。在該階段中，各種類型的相機處理運算例如白平衡，色彩校正，顏色矩陣化，伽瑪校正，位元深度縮減，和壓縮可以用來增強整個圖片品質和/或減少儲存空間，生成了最終的相機輸出 S_d 。 S_d 還可能經過軟體等附加的處理運算，例如，Adobe Photoshop 和 Google Picasa 可以用來進一步提高圖片品質及/或篡改圖像。在圖 1 所示之系統模型中，我們將攝後處理表示為附加的操弄塊，如圖 1 所示。

元件鑑識方法論

正如我們在前一部分系統模型中該討論的那樣，當使用數位相機捕捉真實世界場景時，在最終的數位圖像生成前景景資訊經過了不同的設備元件。該系列資訊處理中的每個



【圖 1】數碼相機中的資訊處理

元件透過採用特定參數集合的特殊演算法對輸入進行修改，在輸出時留下一些固有的指紋痕跡。在接下來的章節中，我們提出利用固有指紋痕跡的非侵入式技術估計各種相機元件之參數。

相機響應函數的估計

相機響應函數 (CRF) 將入射光能量映射為圖像強度值 [3]。CRF 的知識對許多應用都非常有用，例如，利用陰影和光度立體進行銳化的電腦視覺演算法，另外在認證演算法中，可用 CRF 做自然浮水印。基於單個相機輸入的 CRF 估計是一個約束不足問題，因此大部分的前期工作是透過假設一個特定非線性模型估計 CRF。在文獻 [4] 中，Farid 假設 CRF 可以寫成 $f(r) = r^\gamma$ 的形式，其中 r 和 $f(r)$ 分別表示入射光能量和圖像強度值， γ 是轉換參數。Farid 提出，轉換形式 r^γ 在頻域中引入相關性，這可用相干分析計算（三階統計特性），無須瞭解影像裝置的詳情 [4]。儘管此相干分析方法能夠在 7.5% 的平均準確率內估計出 γ 值，該方法受到使用 γ -curve (γ 曲線) CRF 模型的限制，這對於真實世界的 CRF 遠遠不夠。

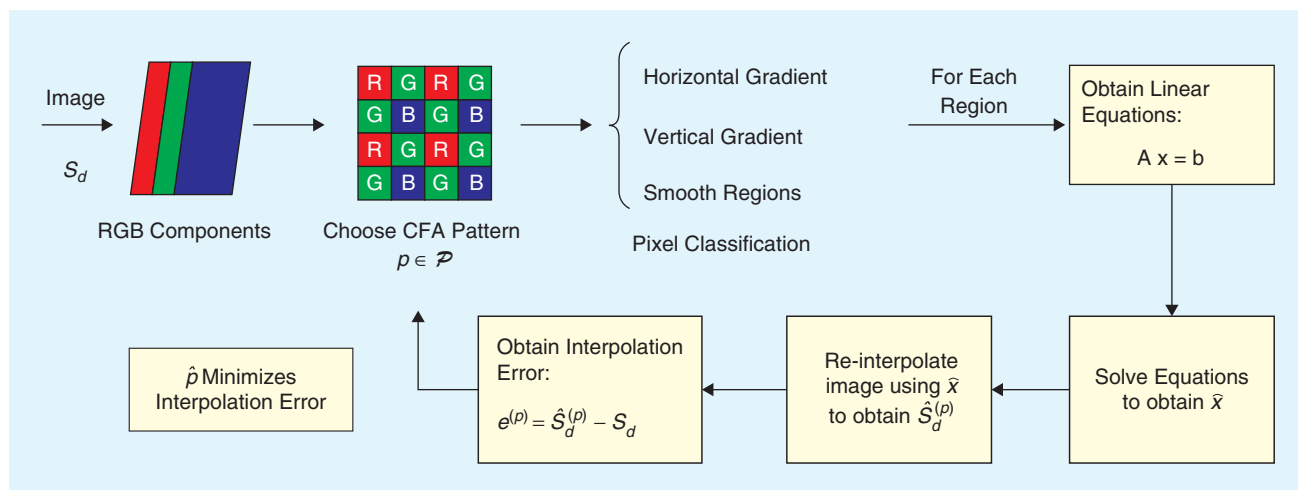
在文獻 [5] 中，透過測算圖像邊緣色彩分佈後非線性響應效果，Lin 和 Zhang 提出了一種從單獨的紅綠藍 (RGB) 顏色圖像估計 CFR 的方法。作者假設邊緣像素是線性混合的，同時引入一種計算逆輻射響應的方法，將邊緣色彩的非線性分佈映射到線性分佈。實驗結果證明，估算的逆回應曲線的平均均方根誤差 (RMSE) 大約為 10^{-2} 。文獻 [6] 將該方法進一步推廣到灰階圖像中，藉由利用沿圖像邊緣的更高階分佈特徵，得到兩個相機裝置的 RMSE 接近 10^{-2} 。Ng 等人在假設 CRF 函數服從一個一階廣義 Gamma 曲線模型 [3]

的條件下，研發了一類有約束方程以識別潛在的局部平面輻照度點，再利用這些點估計 CRF。作者將模擬擴展到五種相機型號，驗證估計演算法之良好性能，該方法用於單幅圖像時，CRF 估計得到的平均 RMSE 接近 10^{-2} ，但 RMSE 隨著額外的相機輸出增加而降低。

彩色濾鏡陣列和色彩插值參數

對於 CFA 和色彩插值模組這些元件而言，元件輸出的知識提供關於輸入的完整資訊，因為輸入和輸出分別對應著採樣和插值資料 [7]。Popescu 和 Farid 採用期望最大 (EM) 演算法來估計鑑識分析中的色彩插值係數。作者先假設圖像像素隸屬於以下兩個模型之一：1) 像素點與其鄰域線性相關，可由一個線性插值演算法獲取，或者 2) 像素點與鄰域不相關。基於此假設，作者提出了兩步 EM 演算法估計 CFA 係數 [7]。在 E 步驟時，估計出每個樣本歸屬於兩個模型的機率，在 M 步驟找出相關性的確定形式。EM 演算法生成兩個輸出：表示像素歸屬於兩個模型可能性的二維機率分佈圖和加權係數。作者透過模擬結果表明，估算的機率分佈圖能有效用於檢測彩色圖像是否經過色彩插值的結果，色彩插值係數則有助於區分不同的插值演算法 [7]。

Swaminathan 等人提出了將估計彩色濾鏡陣列模式和色彩插值係數相結合的演算法 [8]。該方法原理如圖 2 所示。首先基於數碼相機之實際設計建立一個關於 CFA 模式的搜索空間 \mathcal{P} ，同時觀察到大部分採用 RGB 型 CFA 的商業相機所具有的固定週期為 2×2 。對於搜索空間 \mathcal{P} 中的每一個 CFA 模式 p ，在不同類型之紋理區域中，透過線性模型擬合分別計算插值係數。具體而言，基於局部鄰域的梯度特徵，將圖像劃分成三類區域，位置為 (x, y) 的圖像像素被



【圖 2】彩色濾鏡陣列的估計演算法和色彩插值係數

歸類為三個區域之一：區域 \mathcal{R}_1 包含具有顯著水平梯度的圖像部分；區域 \mathcal{R}_2 包含具有顯著垂直梯度的圖像部分， \mathcal{R}_3 包含圖像的剩餘部分，主要包含光滑區域。

利用最終的相機輸出 S_d 和假定的樣本模式 p 直接從感測器陣列擷取的 S_d 中像素位置集合，和被插值之像素位置集合進行識別。對於色彩插值而言，在每個三類區域 $\mathcal{R}_m (m = 1, 2, 3)$ 和圖像的三個彩色通道 (R, G, B) 上假定存在一個線性模型，插值的像素可以表示為像素之加權平均，假定這些像素能直接從感測器擷取。由求解這些方程來獲得係數之權值。假設 N_e 方程組中對於一個特定區域 N_u 未知，色彩通道表示為 $\mathbf{Ax}=\mathbf{b}$ ，其中 \mathbf{A} 的大小是 $N_e \times N_u$ ， \mathbf{b} 的大小是 $N_e \times 1$ ，分別指定可直接擷取的像素值和那些所插值的像素值。 \mathbf{x} 大小是 $N_u \times 1$ ，代表估算的插值係數。為了處理由於插值後的其它機內運算(例如 JPEG 壓縮)而造成的 \mathbf{A} 和 \mathbf{b} 中可能存在的雜訊像素值，用奇異值分解 (SVD) 估計插值係數 [8][9]。得到係數後，用其對相機輸出 S_d 做二次插值，以得到 $\hat{S}_d^{(p)}$ ，並計算誤差項 $e^{(p)} = \hat{S}_d^{(p)} - S_d$ 。對於搜索空間 \mathcal{P} 的所有模式 p ，重複實施這些步驟，插值誤差最低的模式被選為 CFA 的估計。在此過程中，同時獲得與估算的 CFA 模式相對應的插值係數。更多的細節請參閱文獻 [8]。

後插值處理估計

此處理運算是相機在色彩插值後完成的如白平衡和色彩校正等，確保場景中的白色物體在一幅照片中呈現白色。白平衡運算是典型的乘法運算，照片中的每種色彩都乘以一個在相機色彩空間中經適當選擇的常數。由於白平衡運算的乘法性質，非侵入而僅基於單個相機的輸出，無法精確地估計這種運算 [10]。然而，這種運算可以透過非侵入式的兩步方法來估計，首先得到在不同內置白平衡設置下的兩個圖像，然後藉由求解一組方程，利用 Von-Kries 假設推導 [10]。

JPEG 壓縮是數碼相機中流行的另一種後插值處理元件。JPEG 壓縮被看做是在離散餘弦變換 (DCT) 域的量化。在這種情況下，元件輸出的知識不能提供對應於輸入之完整資訊；但有提供量化步長範圍內輸入之粗略估計。文獻 [11] 和 [12] 利用基於像素合併技術統計分析進行量化矩陣的非侵入式估計。已經過論證，這些演算法在圖像的低頻，水平和垂直方向的高頻子帶之量化步長估計方面具有較好的準確性，這些子帶中

非零量化值相當多。在對角方向子帶上，量化為零的係數個數眾多，則導致較大的估計誤差。

元件鑑識之應用

我們現在考慮元件鑑識的一些應用，留在數位圖像上的固有指紋痕跡提供解密線索，有助於解答關於數位圖像之原始性和真實性的許多問題。

相機識別鑑識

估算的相機元件參數將被用作相機識別鑑識的特徵，用來鑒別採集數位圖像的相機品牌和型號。Bayram 等人提出一種相機識別方法 [13]，該方法採用 EM 演算法中的加權係數 [7] 和機率分佈頻譜的尖峰之位置及數量作為特徵。從輸入條件受約束的兩個相機中擷取的圖像，與從網際網路上隨機擷取的圖像一起用於實驗，作者提出三個品牌做實驗時準確率接近 84%，其中 20% 的圖像用於訓練，剩餘的 80% 用於測試 [13]。文獻 [14] 是該方法的進一步改進，分別考慮圖像之光滑和非光滑區域，對於三種相機品牌分類的準確率達到 96%。

Swaminathan 等人在文獻 [8] 中提出了一種結合 CFA 模式和插值演算法的估計技術，用於相機鑑識，針對 19 種不同相機型號，提出對更大資料庫所進行的大量相機識別結果。實驗中包含的相機型號目錄如表 1 所示。對於資料庫中 19 種相機型號的每一種，作者搜集大約 200 幅不同的 512x512 圖像，都是在無約束條件下的不同場景、不同照明環境和不同 JPEG 品質因數的壓縮。根據品牌或型號將資料庫中的這些圖像分為不同的組別，每類區域和色彩通道 (每幅圖像共有 441 個係數) 所估算的 7x7 大小的濾鏡係數被用於相機識別。透過模擬，作者印證了九種不同相機品牌分類的平均識別準確率約為 90%，對於來自九種不同品牌的不同圖像從 19 種相機型號之間區分的識別準確率則接近 86% [8]。此外，作者論證了這些結果不受相機內的後插值處理運算的影響，比如 JPEG 壓縮，加性雜訊和非線性點運算諸如伽瑪 (Gamma) 校正 [8]。

在數量更大和種類更多的資料庫方面，相較於其它關於相機識別的研究，文中的元件鑑識技術具有更高的準確率

[表 1] 文獻[8]的實驗中採用的相機型號

NO.	CAMERA MODEL	NO.	CAMERA MODEL	NO.	CAMERA MODEL
1	CANON POWERSHOT A75	8	NIKON E5400	15	CASIO QV 2000UX
2	CANON POWERSHOT S400	9	SONY CYBERSHOT DSC P7	16	FUJIFILM FINEPIX S3000
3	CANON POWERSHOT S410	10	SONY CYBERSHOT DSC P72	17	FUJIFILM FINEPIX A500
4	CANON POWERSHOT S1 IS	11	OLYMPUS C3100Z/C3020Z	18	KODAK CX6330
5	CANON POWERSHOT G6	12	OLYMPUS C765UZ	19	EPSON PHOTOPC 650
6	CANON EOS DIGITAL REBEL	13	MINOLTA DIMAGE S304		
7	NIKON E4300	14	MINOLTA DIMAGE F100		

[15]–[17]；其後，該技術推廣到照相機上，達到了 98% 的準確率 [35]。Kharrazi 等人在文獻 [15] 中關於相機識別提出了一組 34 個特徵的集合，目的是對數碼相機中的圖像擷取過程建模。特徵集包括：平均像素值、RGB 兩者間的相關性、鄰域分佈重心、RGB 能量比率、小波域統計量 [18]，以及圖像品質度量 [19]。作者採用支撐向量機 (SVM) 演算法進行分類，對在受控輸入條件下來自 3 種品牌 5 種相機型號的測試圖像，公佈其準確率接近 88%。在文獻 [16] 中也在相機識別中使用相同的特徵集，對於在受控輸入條件下來自兩種不同模型的四種相機型號，得到的準確率接近 95%。另一個關於相機識別的研究工作是估算像素之非均勻雜訊，這是圖像非均勻光電響應雜訊之主要分量，是圖像感測器所必然產生的，可用於區別兩個相同品牌，型號和裝置的相機。在演算法的訓練階段，採用一種基於小波去噪演算法取得像素非均勻雜訊估計，該雜訊的隨機分量可藉由平均許多圖像的估計予以消除。在測試階段，判斷某一幅特定圖像是否為某一特定數碼相機所採集，要獲得圖像的雜訊頻譜，並與某特定數碼相機的平均雜訊頻譜（也叫做參考頻譜）求相關。相關值大於所預選的閾值表示某特定圖像來自某特定數碼相機的。作者展示了當採用高品質圖像做測試時，該方法能以 100% 的準確率識別相機之來源。Dirik 等人提出了一種透過粉塵特性識別數碼單反相機 (SLR) 的方法 [20]，對三種不同品牌相機的準確率接近 92%。其針對相機識別的研究工作，重點在於典型特徵的提取，但無法對系列資訊處理的不同元件進行明確的估計。

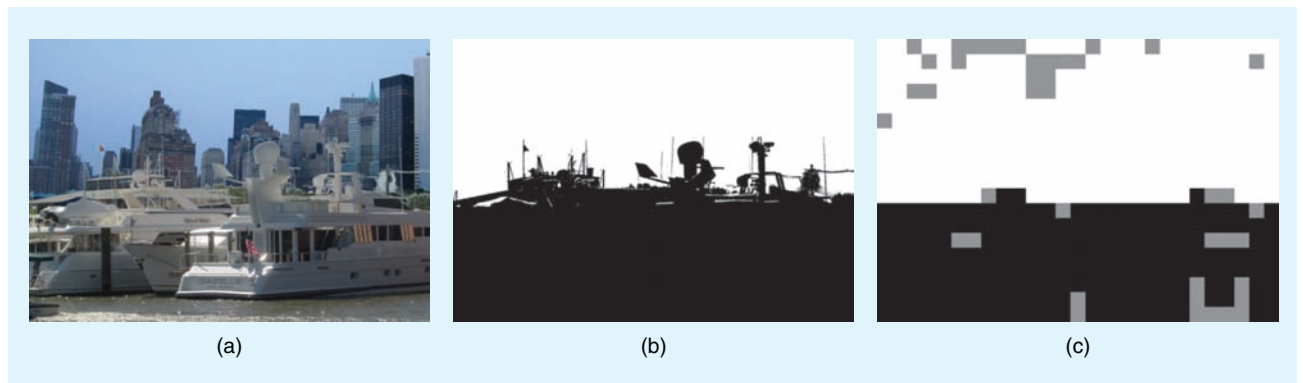
侵權和授權鑑識分析

元件鑑識分析可用來做與影像裝置關聯的共同特徵識別，適合應用於設備元件侵權和授權識別。在文獻 [8] 中，將一種基於分類的方法用於不同相機所使用的插值演算法之

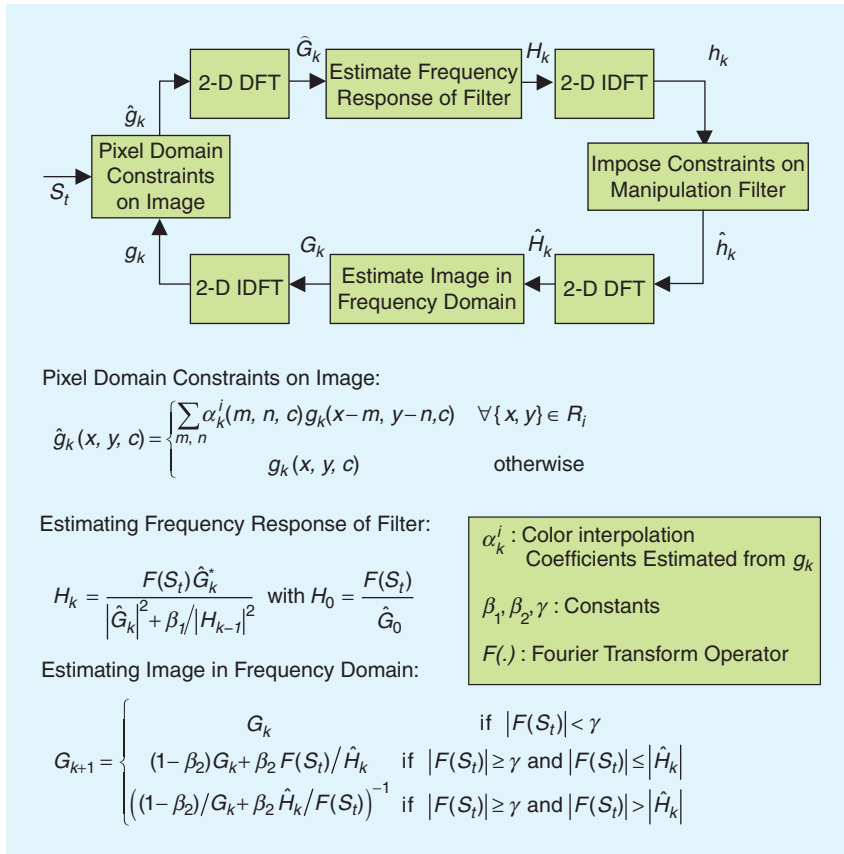
間之相似性研究。作者首先透過從資料中刪掉一種相機型號來訓練分類器，再用刪掉的相機係數對其測試，以找出色彩插值係數空間上的最近鄰域。文章的實驗結果表示，當 SVM 採用來自 18 個相機的 200 幅圖像做訓練（其中不包括 Canon Powershot S410），再使用來自 Canon Powershot S410 的 200 幅圖像做測試時，Canon Powershot S410 圖像的分類準確率為 66%。此外，剩餘圖像的 28% 被分類為 Canon 型號的一種；這意味著不同型號但品牌相同的相機使用這類插值演算法時，存在極大相似性。文獻 [8] 中的結果表明 Minolta DiMage S304 和 Nikon E4300 之間存在相似性，在分類測試時，約為 53% 的 Minolta DiMage S304 圖片被認為是 Nikon E4300 型號。基於這些結果，可定義一種新度量標準來研究兩個相機品牌/型號之間的相似性 [8]。該分析可在估算相機元件參數中的識別相似性方面有所應用，用於確定潛在的侵權或授權。

基於元件參數中不一致性的剪貼鑑識檢測

用剪貼偽造建立的篡改圖像常包含從不同相機擷取的圖片中選取圖像的不同部分，這些相機的內部元件可能採用不同的演算法/參數集合。從不同圖像區域獲得的、估計出來的感測器模式雜訊之間的不一致性 [17]，或者估計出來的相機元件之固有指紋痕跡間的不一致性（例如色彩插值係數 [21] 或 CRFs [22]），可用於識別如剪貼這類的數碼偽造操作。在本文中，我們以文獻 [21] 所提出的研究實例來做論證。在文獻 [21] 中，作者建立一幅大小為 2048x2036 的篡改圖片，這是由使用兩種不同相機所獲得的兩幅圖像經部分組合而得到的。圖 3(a) 和 (b) 分別提出該幅篡改圖像並標記了不同色彩的個別部分。圖 3(b) 中呈現在白色的區域取自於 Canon Powershot S410 數碼相機所拍攝的圖像，黑色部分是 Sony Cybershot DSC P72 型相機所攝的圖片裁



【圖 3】來源鑒定之應用 (a) 篡改的圖像樣本，(b) 從兩個相機擷取的區域 (c) CFA插值識別結果(黑色：Sony Cybershot DSC P72；白色：Canon Powershot S410；灰色：分類為其它相機的區域)



[圖 4] 用以估算操弄濾波器係數的反覆運算約束增強演算法 [24]

剪和粘貼而成。合成的圖像經過了品質因數為 80% 的 JPEG 壓縮。

為了在圖片的不同部分上識別其固有的相機指紋，利用一個步階為 64x64，大小為 256x256 的滑窗進行圖像測試，估計每一個 256x256 塊內的色彩插值係數 [21]。19 個相機型號分類器的檢測結果如圖 3(c) 所示。圖中標記為黑色的區域表示那些被分類為 Sony Cybershot DSC P72 機型，而白色區域則相對應來那些被正確分類為 Canon Poweshot S410 機型的部分。用灰色表示的剩餘區域對應於那些被錯誤分類為剩下 17 個機型中一種的圖塊。正如圖 3(c) 所示，篡改圖像的大部分區域內，利用 256x256 大小的巨集區塊擷取的資料可以識別出正確的相機，並且具有較高可信度。

[表 2] 實驗中包含的篡改運算

MANIPULATION OPERATION	PARAMETERS OPERATION	OF	THE NUMBER OF IMAGES
SPATIAL AVERAGING	FILTER ORDERS 3-11 IN STEPS OF TWO		5
MEDIAN FILTERING	FILTER ORDERS {3, 5, 7}		3
ROTATION	DEGREES {5, 10, 15, 20}		4
RESAMPLING	SCALE FACTORS {0.5, 0.7, 0.85, 1.15, 1.3, 1.5}		6
ADDITIVE NOISE	PSNR 5 DB AND 10 DB		2
HISTOGRAM EQUALIZATION			1
TOTAL			21

在上述特例中，被操弄圖像具有來自兩個不同相機的明顯痕跡，因此是被篡改的。

用作篡改檢測中真值建模的元件鑑識

攝後處理運算包括內容保留和內容修改操弄，例如篡改。攝後處理通常難以檢測和估計，由於缺乏關於操弄種類的知識，這導致對模型的不當選擇。為了避免此問題，一些關於篡改檢測的早期文獻試圖透過從經歷的失真這種形式來定義操弄後的圖像特性以檢測篡改，利用該分析提出了用於檢測操弄後的圖像和識別操弄種類及其參數。在文獻 [18] 和 [23] 中，採用基於變異數分析方法 [23] 和高階小波統計量 [18] 的特徵檢測圖像的操弄是否存在，但沒有關注於識別操弄種類和/或其係數等方面。這些方法用於分類時需要篡改後的圖像樣本（針對每種操弄），從真正的相機採集圖像中區分這些操弄過的圖像。另外，這些圖像並不能有效識別操弄的種類，因為構

建分類器時並沒有對其建模或直接考慮在內。

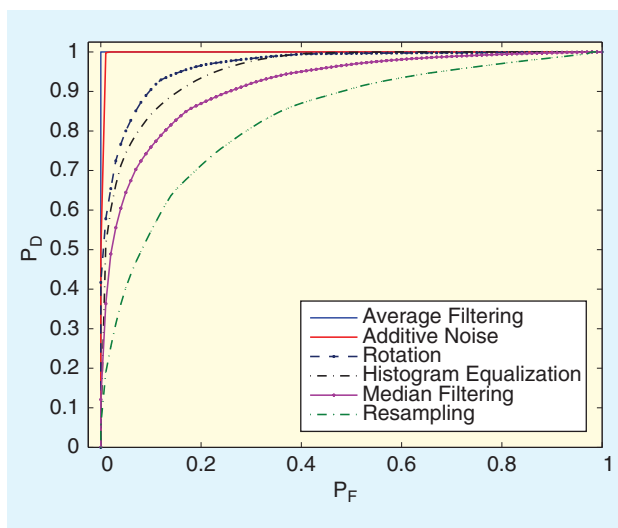
在文獻 [24] 中，作者透過把所有攝後處理作為一個操弄塊進行建模，將元件鑑識方法推廣到用於識別圖像操弄的存在與否 [24]。該演算法首先假設某特定測試圖像， S_t 是操弄過的相機輸出，由實際的相機輸出 S_d 經過一種操弄運算得到。任何應用在 S_d 上的後相處理這時都可作為某種線性濾波器進行建模，其係數可用反覆運算的約束增強演算法估計 [24]。圖 4 提出了文獻 [24] 中用來估計操弄濾波器的係數之演算法原理圖。測試圖像 S_t 則用於反覆運算過程的初始化。對每一反覆運算，對圖像和濾波器在像素域及傅立葉域中反覆地運用已知的約束條件，以更新所估計的相機輸出 g 和濾波係數 h [25] [26]。在第 k 次反覆運算時，對圖像

g_k 施加像素域約束條件以得到 \hat{g}_k ；像素域約束條件代表相機約束條件，這裡相機元件參數 α_k^i 可以透過章節“彩色濾鏡陣列和色彩插值參數的估計”中的元件鑑識技術估計。在獲得圖像 \hat{g}_k 後，對其做

離散傅立葉變換 (DFT) 以得到 G_k 。所估計的操弄濾波器之頻率回應可由 \hat{G}_k 和測試圖像的傅立葉變換 $\mathcal{F}(S_i)$ 而得，記為 H_k ，如圖 4 所示。所估計的回應 H_k 進行反傅立葉變換可得到 h_k ，為獲得 \hat{h}_k 而施加給 h_k 的濾波器約束條件是 h_k 的實部存在。最終獲得的 G_{k+1} 值是一個雙變數估計的函數形式：a) 當前的 G_k 值和 b) $\mathcal{F}(S_i)/\hat{H}_k$ ，其中 $\hat{H}_k = \mathcal{F}(\hat{h}_k)$ 。該演算法的完整細節及其性質參見文獻 [24]。所估計出的操弄濾波器參數與恆等變換間的相似性可用以表明在相機擷取後測試圖像是否經過某些操弄。

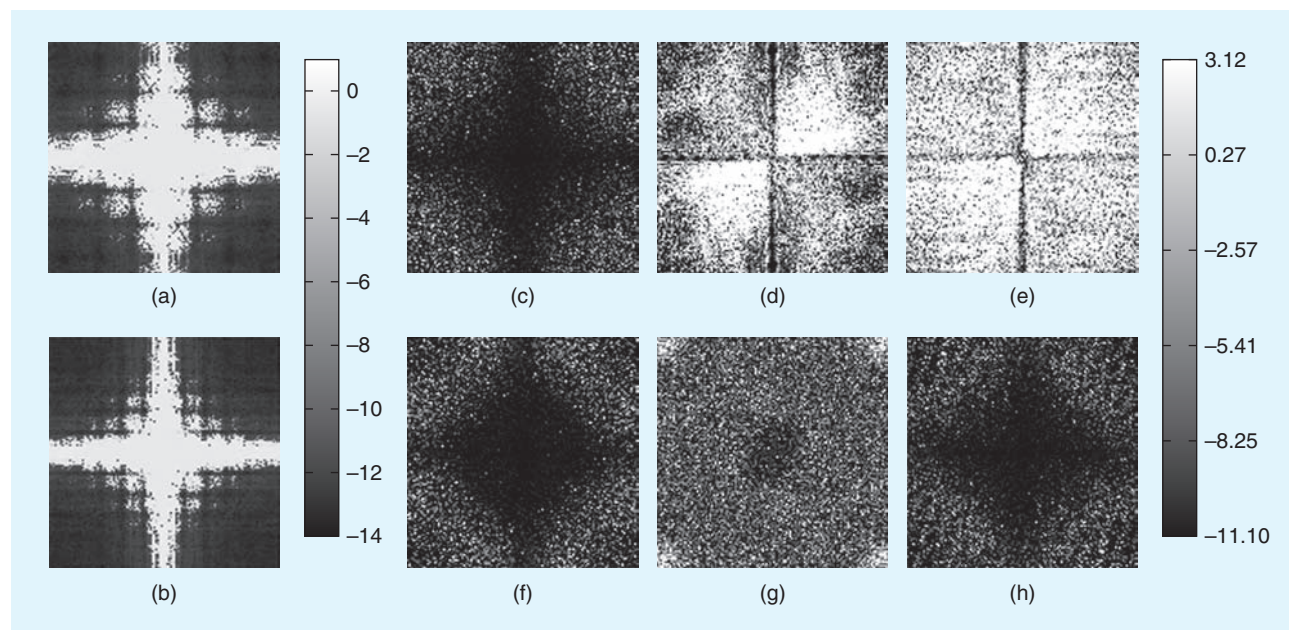
在文獻 [24] 中，作者採用來自 9 種不同相機型號的資料 (對應的機型編號是表 1 中的 1-7、10 和 16) 用於測試反覆運算的約束增強演算法。這裡提出了在相機圖像庫中 900 種不同的 512x512 圖像，每個相機型號具有 100 個圖像。這些圖像經過處理加工，每個圖像生成 21 種篡改版本，共獲得 18,900 幅操弄過的圖像。表 2 中列出了考察的操弄種類及其參數設置。對於每一種直接的相機輸出及其操弄過的版本，計算估計的操弄濾波器的頻域係數 H_i ，與選擇的參考型號 H_{ref} 之間的相似度，使用一種相似性值來定義。真實的相機輸出的參考型號 H_{ref} ，在訓練階段可利用同樣的反覆運算約束增強演算法優先獲得，有助於抵消相機內的後插值處理產生的較小誤差。為了計算相似性值，得到的測試圖像頻率回應幅值的對數函數 $\Theta_i = \log_{10}(|H_i|)$ ，與參考圖像的對數幅度係數之間相似性計算如下：

$$s(\Theta_i, \Theta_{ref}) = \sum_{m,n} (\Theta_i(m,n) - \mu_i) \times (\Theta_{ref}(m,n) - \mu_{ref})$$



【圖 5】關於篡改檢測的接受者操作特徵 (ROC)，這時測試物件是資料庫中所有圖像，其中 200 幅圖像用來訓練。

這裡 μ_i 表示 Θ_i 的像素均值， μ_{ref} 表示 Θ_{ref} 的像素均值。假如與參考型號的相似性大於適當選定的閾值，則測試的輸入則被分類為未經操弄的。另一方面，假如輸入圖像經過了篡改或加密寫入運算，估算的操弄濾波器係數應包含這些操弄的影響，因此與參考型號 (從一個未經操弄的相機輸出得到) 之間的相似度會減少，這樣會導致相似性值將低於選定的閾值。圖 5 以接受者操作特徵 (ROC) 的形式提出了反覆運算條件增強演算法的性能。圖中顯示，在 P_F 接近 10% 的情況下，對於諸如空域求平均和加性雜訊



【圖 6】對相機輸出的操弄濾波器頻率回應 (a) 7x7 均值濾波器，(b) 11x11 均值濾波器，(c) 7x7 中值濾波器，(d) 20 度旋轉，(e) 70% 重採樣，(f) 130% 重採樣，(g) 雜訊加至 PSNR20dB，(h) 長條圖均衡化。頻率回應在 log 尺度上展示和移動，以便 DC 分量處於中心。

等操弄， P_D 接近 100%，對於中值濾波，長條圖均衡化和旋轉則約為 70%~80%。

在文獻 [24] 中，作者利用估算的操弄濾波係數識別攝後處理運算的種類和係數。圖 6 提出了對於表 2 中不同種類的操弄，所估計的操弄濾波係數之頻率回應。仔細觀察頻域中的操弄濾波係數，可看出不同類型的篡改操弄間有明顯區別。對於一些如均值濾波的操弄，可在頻譜上觀察到明顯的空白，這些空白的間隙可以用來估計均值濾波的階數及其參數。例如加性雜訊這樣的圖像操弄會導致在操弄濾波器中具有一個雜訊譜如圖 6(g) 所示，雜訊強度可由濾波係數計算而來。旋轉和向下採樣可以由操弄濾波器頻譜 LH 和 HL 子帶上較小的值識別。近年來，Chuang 等人 [27] 在文獻 [24] 基礎上提出了許多種類的線性移不變 (LSI) 和非 LSI 影像處理運算，例如重採樣，JPEG 壓縮，和非線性濾波，在它們的經驗頻率響應 (EFR) 中展現了一致和獨特的圖譜。在用做六類操弄分類時，對基於 EFR 的操弄種類，識別性能大約為 93%。

通用隱藏分析和圖像擷取鑑識之應用

就通用隱藏分析和圖像擷取鑑識分析而言，其共同挑戰就是如何對真實的原始圖像資料進行建模。使用一個相機型號及其元件分析，元件鑑識可以提供一個框架來區分相機擷取圖像和具有加密隱藏資訊的圖像或其它方式擷取的圖像 [24]。圖像插值，例如浮水印和加密，可被建模為一種後處理操弄並應用於相機輸出，由反覆運算約束增強演算法估計出的操弄係數可用來將其與原始資料做區分。由其它類型之擷取來源生成的圖像會導致操弄濾波器與期待得到相機輸出的理想 δ 函數明顯不同，該圖像可藉由對比操弄濾波係數和從相機輸出直接得到的參考型號來區分。其它細節請見文獻 [24]。

元件鑑識的理論分析

正如先前章節所述，經過影像裝置的不同元件後，其固有指紋痕跡留在最終的數碼照片上，可用作證據來估計元件參數並提供線索，以解答關於數位資料之原始性和真實性等鑑識問題。然而，由於固有指紋痕跡經過系列資訊處理之不同部分，因此其中的一些可能已經改變或者損毀，而另外一些又重新生成。這引發了許多基本問題，例如什麼元件痕跡丟失或者改變了？哪些元件在系列資訊處理中可識別而哪些又不是？一個元件的可識別性會如何影響另一個的估計？關於元件鑑識之理論框架在文獻 [28] 和 [29] 中提出，它們關注於解答此類問題及測試元件參數能被識別

或者精確分類之條件。下面，我們將總結該工作的主要結果。

理論概念和框架

在文獻 [28] 和 [29] 中，作者定義了元件做為系列資訊處理之基本單元，將設備表示為 N_c 個元件的串聯，寫成 $\{C_1, C_2, \dots, C_{N_c}\}$ 的形式。作者將系統的第 k 個元件 C_k 所採用的參數集合記為 θ^k ，元件鑑識分析之目標即根據元件參數 θ^k 的可識別性或者可分類來定量。關於 θ^k 作者考慮了兩種可能情境。在第一種情況下，作者假定演算法空間關於 θ^k 之可能集合為事前已知，也就是說 $\theta^k \in \Theta^k = \{\theta_1^k, \theta_2^k, \dots\}$ 。在這種情況下，此問題轉化為一個分類問題，可以用模式分類方法分析該情境 [28]。在第二種情境下，假設關於可能演算法空間的知識並非事前已知，或者不是假定的事前機率，作者提出了一個基於估計理論和 Fisher 資訊的框架分析此情況 [29]。

若 θ^k 取機率有限的值，這樣 $\theta^k \in \Theta^k$ ，對於該情境，作者定義元件 C_k 為侵入可分類 (i-classifiable) 的條件為：對所有輸入，利用正確演算法進行分類可侵入分類或者說是 i-classifiable，把元件分類到利用正確演算法，在某特定輸入和輸出的情況下，其機率大於或等於將其錯分到其它類別之機率。此外，對於至少一個輸入 x^* 及其相應輸出，正確分類之機率完全大於錯分機率 [28][30]。在這種情境下，參數識別演算法的優勢是在作出正確判斷時可以根據置信度衡量，可被定義為做出正確判斷之可能性和做出錯誤判斷相應之最大可能性之差。如此這般，對於正確判斷 θ^k 的置信度 $\gamma_i^k(x)$ 可以寫為：

$$\gamma_i^k(x) = f(\theta_i^k | y, x) - \max_{\theta \in \Theta^k \setminus \theta_i^k} f(\theta | y, x)$$

這裡 $f(\theta | y, x)$ 表示元件在輸入為 x ，相應的輸出為 y 之條件下，元件採用參數 θ 的機率

置信度 $\gamma_i^k(x)$ 是一個輸入為 x 的函數，可以透過選擇合適的輸入將其提升。例如，考慮一個具有參數 $\{\xi_0, \xi_1\}$ 的元件做例子，其輸入和輸出由以下關係提出：

$$y(n) = \xi_0 x(n) + \xi_1 x(n-1)$$

設 $x^{(1)} = [\dots, 1, 1, 1, \dots]$ ， $x^{(2)} = [\dots, 0, 1, 2, \dots]$ 是系統的兩個可能輸入。相對應的輸出分別是 $y^{(1)} = [\dots, \xi_0 + \xi_1, \xi_0 + \xi_1, \xi_0 + \xi_1, \dots]$

和 $y^{(2)} = [\dots, -\xi_1, \xi_0, 2\xi_0 + \xi_1, \dots]$ 。請注意， $y^{(1)}$ 是一個常數序列，每個元素都是等於 $\xi_0 + \xi_1$ 的，該和式不能提供關於參數 ξ_0 或 ξ_1 的任何提示。因此，對於元件值的估計而言， $x^{(1)}$ 並不是較好的輸入。另一個方面，觀察

系統的輸出 $y^{(2)}$ ，可以表示為一個線性方程的系統以計算 ξ_0 或 ξ_1 值，因此， $x^{(2)}$ 是可以得到元件參數值的較好輸入。本例說明了參數估計中的置信度可以透過輸入的選擇來提高，將該結論推廣，文獻 [28] 定義了使置信度達到的最佳輸入。

在文獻 [28] 中，半非侵入式可分和完全非侵入式可分元件的定義相似，這些定義有利於確立一些理論成果。例如，作者已證明假如一個元件是非侵入式可分的，然而其參數可以被半非侵入式識別，假如一個系統是半非侵入式可分的，那麼其每個元件也是侵入式可分的。另外，利用半非侵入式分析得到的平均置信度值都高於或等於那些透過完全非侵入式分析得到的結果，低於那些透過侵入式分析所得的結果。此結果是因為半非侵入式鑑識對鑑識分析提供更多的控制，因此可以設計更好的輸入以提高整體性能，侵入式分析在實驗裝置上提供最高的控制。理論上的結果也已驗證當且僅當系統的所有元件都一致時，侵入式，半非侵入式和完全非侵入式鑑識可以提供相同的置信度，這意味著輸入的知識可以提供關於輸出的全部資訊，反之亦然[28]。

圖像採集技術的知識有助於進一步解答相關的鑑識問題，特別是圖像在成像之後後可能經過了怎樣的其他操弄

在第二個情境下，如果關於可能演算法空間的先備知識不存在，那麼元件鑑識問題就變成一個估計問題，估計偏差和估計變異數作為理論上的分析度量，這部分內容在文獻 [29] 中有所討論。對於這類情況，透過半非侵入式分析得到的元件參數估計誤差小於透過完全非侵入式分析所得的結果，而大於侵入式分析所得結果，對於一致元件，這些分析技術是等價的。定義和原理的細節及證明示意圖如文獻 [29] 所示。

數碼相機之實例研究

我們現在考察一些實例研究以論證理論框架的實用性

■ 彩色濾鏡陣列和色彩內插模組：在存在雜訊或者附加的處理時，對於諸如 CFA 和色彩內插模組這類元件，元件輸出的知識提出了關於對應輸入的完整資訊，因為輸入和輸出分別對應著經過採樣和插值的資料。因此，在該情境下，CFA 和色彩內插模組皆為一致的元件。前述之理論分析已經證明半非侵入式鑑識可以提供與完全非侵入式鑑識同樣的準確率，也就是說，即使在輸入條件受約束，以及輸入需精心設計的情況下，元件估計的準確率不能提高到與非侵入式分析相提並論。

然而，存在附加的後插值處理運算時，元件就不再一致，半非侵入式鑑識可以提供優於完全非侵入式鑑識的準確性[28][29]。該情況下，有效的測試條件和輸入圖像已經在基於這類元件的背景知識上被設計出來 [10]，並且在 [28] [29] 被進一步優化以達到更高的置信度和準確性。

■ 後插值處理模組：諸如白平衡和色彩修正等運算，實際上都是典型的乘法運算。由於其乘法特性，皆不是非侵入式可分的 [10]，例如某個特定兩項乘積形式的輸出，那麼不足以把單個的項式分別地明確解析。在該情境下，作者證實關於相機輸入內容的資訊有助於處理該問題，半非侵入式分析被用以參數估計並具有良好的準確性 [10]。

■ 後-相機處理模組：許多後-相機處理模組可以透過理論框架進行類似地分析。為了實施非侵入式估計，一些文獻提出了這種後-處理運算的方法，例如重採樣 [31]，不

規則雜訊頻譜 [32]，亮度或照明方向 [33]，色差 [34]，非線性點運算，和伽瑪 (Gamma) 校正 [4]。例如，當圖像上採

樣時，一些像素點值從較小的圖像直接獲得，留下的像素則藉由插值擷取，因此與其鄰域高度相關。這樣，重採樣參數可以由對特定範圍內重採樣像素值引入的相關性研究來識別 [31]。諸如對比度變化，伽瑪 (Gamma) 校正和其它圖像非線性化等影像處理可以被建模，更高階的統計例如雙頻譜可以被用於識別其元件參數 [32]。其中的一部分方法假定關於可能演算法空間的先備知識，需要對所有機率進行徹底搜索。

本章節考察的理論分析框架可以提供系統方法論用於解答系列資訊處理中哪些元件和處理運算是可識別的或者是不可識別的，這將有助於衡量估計之準確性。這些框架可被推廣到一系列廣義資訊處理中對不同元件間交互作用的研究上。

結論

本文考慮了元件鑑識問題，提出關於視覺感知之多媒體元件鑑識的當前文獻資料調查。本文安排為三個部分。在第一部分中討論了許多數碼相機元件鑑識之方法論，這些方法可以用於估計相機內元件，例如相機響應函數，彩色濾鏡陣列和色彩內插參數，另外也討論了後插值處理演算法，例如白平衡和 JPEG 壓縮。第二部分展示了可以用於

多種不同應用之估計參數，包括設備品牌和型號識別，侵權/認證鑑識分析，建立真值模型以檢測全域和局部的篡改，包括加密寫入，對圖像採集鑑識進行不同擷取來源圖像的區分。第三部分提出了關於元件鑑識之理論分析框架，主要關注在得到關於元件鑑識的具體理解，和解答許多關於何種處理運算可以或不可以將其識別，以及在什麼條件下可以識別這類基本問題。總的來說，我們認為元件鑑識分析對於專利侵權案例，智慧財產權管理和數位傳媒之技術進化研究等，可以提供一個資訊的主要來源，並且推動多媒體鑑識發展，以便深入理解資訊處理鏈。

作者簡介

Ashwin Swaminathan (sashwin@qualcomm.com) received the B.Tech degree in electrical engineering from the Indian Institute of Technology, Madras, India in 2003, and the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park in 2008. He is currently a senior engineer at Qualcomm Incorporated in San Diego, California. He was a research intern with Hewlett-Packard Labs in 2006 and Microsoft Research in 2007. His research interests include multimedia forensics, information security, authentication, and information discovery. He was the winner of the Student Paper Contest at the 2005 IEEE International Conference on Acoustic, Speech and Signal Processing and received the ECE Distinguished Dissertation Fellowship Award in 2008. He is a Member of the IEEE.

Min Wu (minwu@eng.umd.edu) received the Ph.D. degree in electrical engineering from Princeton University in 2001. She is an associate professor at the University of Maryland, College Park. Dr. Wu leads the Media and Security Team at the University of Maryland, with main research interests on information security and forensics and multimedia signal processing. She is a corecipient of two Best Paper Awards from the IEEE Signal Processing Society and EURASIP, respectively. She also received an U.S. NSF CAREER award, a TR100 Young Innovator Award from the *MIT Technology Review Magazine*, a U.S. ONR Young Investigator Award, and a Computer World “40 Under 40” IT Innovator Award. She is currently the area editor of *IEEE Signal Processing Magazine* for its “Inside Signal Processing E-Newsletter” and is the associate editor of the *IEEE Transactions on Information Forensics and Security*.

K.J. Ray Liu (kjrliu@eng.umd.edu) is a Distinguished Scholar-

Teacher of University of Maryland, College Park, where he received university-level Invention of the Year Award; and both Poole and Kent Senior Faculty Teaching Award and Outstanding Faculty Research Award from A. James Clark School of Engineering Faculty. Dr. Liu is the recipient of numerous best paper awards and was an IEEE Signal Processing Society Distinguished Lecturer. He was Vice President- Publications, the Editor-in-Chief of *IEEE Signal Processing Magazine*, and the founding Editor-in-Chief of *EURASIP Journal on Applied Signal Processing*.

參考文獻

- [1] J. Adams, K. Parulski, and K. Spaulding, “Color processing in digital cameras,” *IEEE Micro*, vol. 18, no. 6, pp. 20–30, Nov./Dec. 1998.
- [2] J. E. Adams, “Interaction between color plane interpolation and other image processing functions in electronic photography,” in *Proc. SPIE Cameras and Systems for Electronic Photography & Scientific Imaging*, San Jose, CA, Feb. 1995, vol. 2416, pp. 144–151.
- [3] T.-T. Ng, S.-F. Chang, and M.-P. Tsui, “Using geometric invariants for camera response function estimation,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Minneapolis, MN, June 2007, pp. 1–8.
- [4] H. Farid, “Blind inverse gamma correction,” *IEEE Trans. Image Processing*, vol. 10, no. 10, pp. 1428–1433, Oct. 2001.
- [5] S. Lin, J. Gu, S. Yamazaki, and H.-Y. Shum, “Radiometric calibration from a single image,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Washington, D.C., June 2004, vol. 2, pp. 938–945.
- [6] S. Lin and L. Zhang, “Determining the radiometric response function from a single grayscale image,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, San Diego, CA, June 2005, vol. 2, pp. 66–73.
- [7] A. C. Popescu and H. Farid, “Exposing digital forgeries in color filter array interpolated images,” *IEEE Trans. Signal Processing*, vol. 53, no. 10, part 2, pp. 3948–3959, Oct. 2005.
- [8] A. Swaminathan, M. Wu, and K. J. R. Liu, “Non-intrusive component forensics of visual sensors using output images,” *IEEE Trans. Inform. Forensics Sec.*, vol. 2, no. 1, pp. 91–106, Mar. 2007.
- [9] C. F. van Loan, *Introduction to Scientific Computing: A Matrix-vector Approach Using MATLAB*. Englewood Cliffs, NJ: Prentice-Hall, 1999.
- [10] A. Swaminathan, M. Wu, and K. J. R. Liu, “Optimization of input pattern for semi non-intrusive component forensics of digital cameras,” in *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing*, Honolulu, HI, Apr. 2007, vol. 2, pp. 225–228.
- [11] J. Lukas and J. Fridrich, “Estimation of primary quantization matrix in double compressed JPEG images,” in *Proc. Digital Forensics Research Workshop*, Cleveland, OH, Aug. 2003.
- [12] Z. Fan and R. L. de Queiroz, “Identification of bitmap compression history: JPEG detection and quantizer estimation,” *IEEE Trans. Image Processing*, vol. 12, no. 2, pp. 230–235, Feb. 2003.
- [13] S. Bayram, H. T. Sencar, N. Memon, and I. Avcibas, “Source camera identification based on CFA interpolation,” in *Proc. IEEE Int. Conf. Image Processing*, Genoa, Italy, Sept. 2005, vol. 3, pp. 69–72.
- [14] S. Bayram, H. T. Sencar, and N. Memon, “Improvements on source camera-model identification based on CFA interpolation,” in *Proc. WG 11.9 Int. Conf. Digital Forensics*, Orlando, FL, Jan. 2006.
- [15] M. Kharrazi, H. T. Sencar, and N. Memon, “Blind source camera identification,” in *Proc. Int. Conf. Image Processing*, Singapore, Oct. 2004, vol. 1, pp. 709–712.

- [16] M.-J. Tsai and G.-H. Wu, "Using image features to identify camera sources," in *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing*, Toulouse, France, May 2006, vol. 2, pp. 297–300.
- [17] M. Chen, J. Fridrich, M. Goljan, and J. Lukas, "Determining image origin and integrity using sensor pattern noise," *IEEE Trans. Inform. Forensics Sec.*, vol. 3, no. 1, pp. 74–90, Mar. 2008.
- [18] H. Farid and S. Lyu, "Higher-order wavelet statistics and their application to digital forensics," in *IEEE Workshop on Statistical Analysis in Computer Vision*, Madison, WI, June 2003, vol. 8., pp. 94–101.
- [19] I. Avciabas, B. Sankur, and K. Sayood, "Statistical evaluation of image quality metrics," *J. Electron. Imag.*, vol. 11, no. 2, pp. 206–223, Apr. 2002.
- [20] E. A. Dirik, H. T. Sencar, and N. Memon, "Source camera identification based on sensor dust characteristics," in *Proc. IEEE Workshop Signal Processing Applications for Public Security and Forensics*, Brooklyn, NY, Apr. 2007, pp. 1–6.
- [21] A. Swaminathan, M. Wu, and K. J. R. Liu, "Component forensics of digital cameras: A non-intrusive approach," in *Proc. Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 2006, pp. 1194–1199.
- [22] Y.-F. Hsu and S.-F. Chang, "Image splicing detection using camera response function consistency and automatic segmentation," in *Proc. IEEE Int. Conf. Multimedia and Expo*, Beijing, China, July 2007, pp. 28–31.
- [23] I. Avciabas, S. Bayram, N. Memon, M. Ramkumar, and B. Sankur, "A classifier design for detecting image manipulations," in *Proc. Int. Conf. Image Processing*, Singapore, Oct. 2004, vol. 4, pp. 2645–2648.
- [24] A. Swaminathan, M. Wu, and K. J. R. Liu, "Digital image forensics via intrinsic fingerprints," *IEEE Trans. Inform. Forensics Sec.*, vol. 3, no. 1, pp. 101–117, Mar. 2008.
- [25] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Mag.*, vol. 13, no. 3, pp. 43–64, May 1996.
- [26] G. R. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its applications," *Opt. Lett.*, vol. 13, no. 7, pp. 547–549, July 1988.
- [27] W.-H. Chuang, A. Swaminathan, and M. Wu, "Tampering identification using empirical frequency response," in *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing*, Taipei, Taiwan, Apr. 2009.
- [28] A. Swaminathan, M. Wu, and K. J. R. Liu, "A pattern classification framework for theoretical analysis of component forensics," in *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing*, Las Vegas, NV, Apr. 2008, pp. 1665–1668.
- [29] A. Swaminathan, M. Wu, and K. J. R. Liu, "A component estimation framework for information forensics," in *IEEE Workshop on Multimedia Signal Processing*, Crete, Greece, Oct. 2007, pp. 397–400.
- [30] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley-Interscience, 2000.
- [31] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling," *IEEE Trans. Signal Processing*, vol. 53, no. 2, pp. 758–767, Feb. 2005.
- [32] A. C. Popescu and H. Farid, "Statistical tools for digital forensics," in *Proc. 6th Int. Workshop Information Hiding & Lecture Notes in Computer Science*, Toronto, Canada, May 2004, vol. 3200, pp. 128–147.
- [33] M. K. Johnson and H. Farid, "Exposing digital forgeries in complex lighting environments," *IEEE Trans. Inform. Forensics Sec.*, vol. 2, no. 3, part 1, pp. 450–461, Sept. 2007.
- [34] M. K. Johnson and H. Farid, "Exposing digital forgeries through chromatic aberration," in *Proc. ACM Multimedia and Security Workshop*, Geneva, Switzerland, Sept. 2006, pp. 48–55.
- [35] C. E. McKay, A. Swaminathan, H. Gou, and M. Wu, "Image acquisition forensics: Forensic analysis to identify imaging source," in *Proc. IEEE Conf. Acoustic, Speech, and Signal Processing*, Las Vegas, NV, Apr. 2008, pp. 1657–1660.

[SP]

機率主題模型

聚焦文件和圖像分析之圖形模型設計和應用

本

文中，我們將回顧機率主題模型：可透過數量較少的字詞分佈概括文件集合之圖形模型。稱這些分佈為「主題」，皆因當其與資料相契合時，即可捕捉到貫穿整個群集的突出主題。我們描述了兩個有限維參數化主題模型及其對應的基於階層 Dirichlet 過程 (HDP) 之非參數貝氏方法。討論主題模型對對的時間序列資料的兩種擴展，其中之一假設主題隨著時間的推移緩慢變化，另一個則假定主題變化之發生率。最後，我們將闡述主題模型對非文本資料之應用，並總結其在影像分析中最近的一些研究成果。

前言

階層混合建模已經成為在大量雜亂資料中尋找模式和結構之強大方法。階層混合模型最近頗為成功的應用即為主題建模，其中需要研究的資料是大量文件，混合建模演算法



可以找到嵌於文件集合的字詞基礎模式（參見圖 1）[15] [35]。對大量語料文件實施有效聚類、搜尋、排序、探索、預測和總結，發現這些模式，即是主題。

主題建模演算法通過使用階層混合模型計算，是一種快速演算法，儘管是為分析文件而開發的一種手段，目前已成功應用於多個領域。例如，將主題建模演算法用於尋找影像、音樂 [34]、音訊和語音 [33] [49]、基因

資料 [46]、電腦程式碼 [4]，甚至建築發掘 [40] 的模式。特別是廣泛應用的電腦視覺 [6] [10] [16] [19] [28] [37] [57] [59]，研究人員已在各種電腦視覺問題中使用主題模型。範例包括排序多個影像至不同的場景層次類，用字詞標註影像，分割和標貼影像內的物件。最近，研究人員已擴展至諸如視訊分析的統計主題模型 [60] [61]。

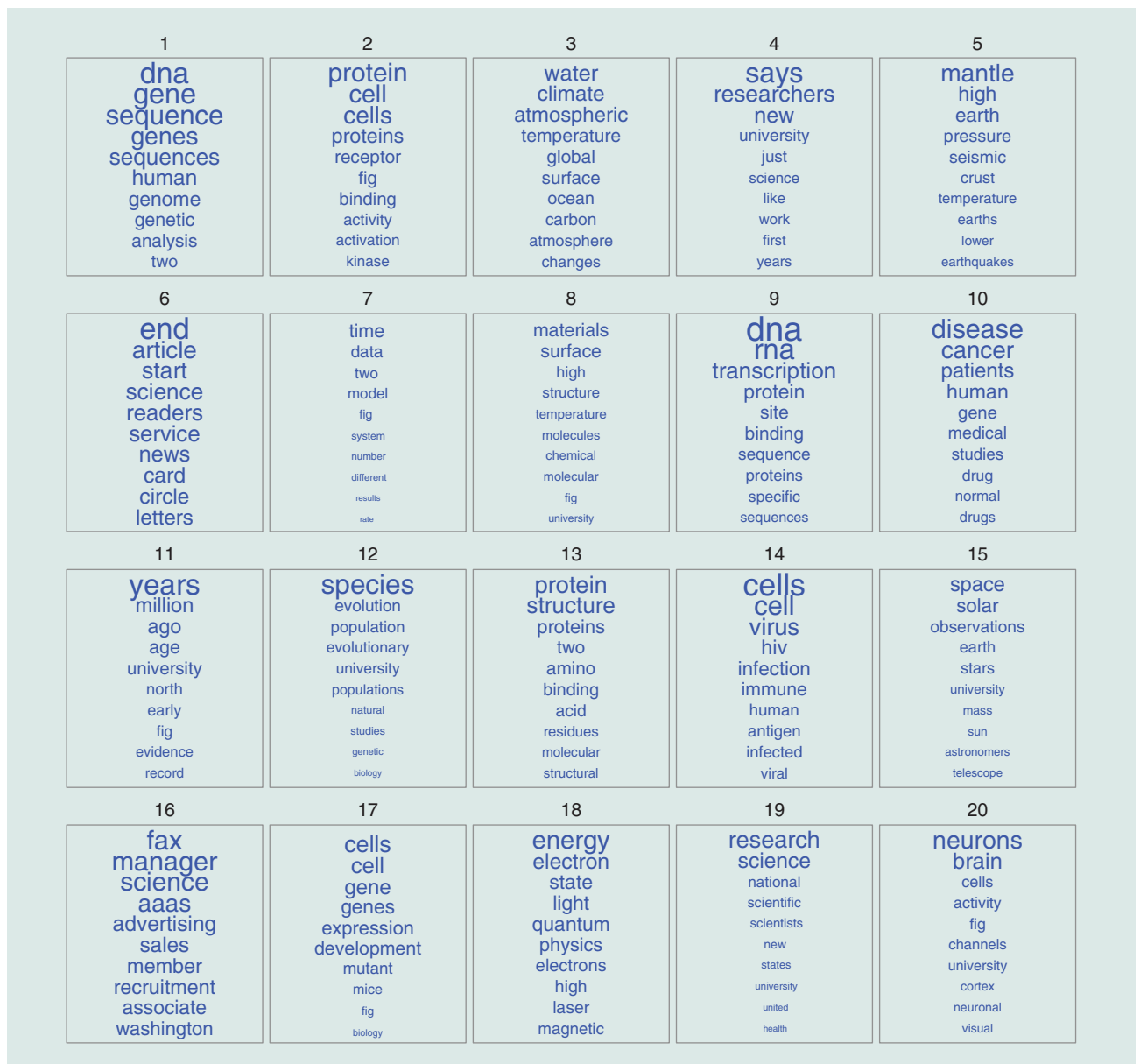
此外，主題建模之新應用推動階層混合模型之新的統計發展。原本主題模型假定資料可交換，亦即集中文件的順序並不重要，文件中單詞的順序亦不重要。於許多問題而言過於嚴苛，放寬這一假設是建立更好主題模型之重要方式；主題模式已在空間設置 [3] [24] [53]、時間序列設置

[14] [47] [62]，和依賴於外在共變數的設置 [39] 中得到應用和延續。

主題建模中另一個值得注意的統計發展是將其推廣至非參數貝氏方法。原始主題模型是有限階層混合，亦即參數化混合模型。非參數貝氏建模之最新進展，特別是 HDP [54]，已經導向「無限」之主題模型。主題的數量無需事先指定，並能隨著集合的擴增而增長。

在這篇文章中，我們將回顧階層混合模型的最新進展，並特別將重點放在主題模型。首先關注分析文件，敘述有

限主題模型和貝氏非參數之主題模式（無限容量）。然後，我們將敘述循序分析集合的延伸，以說明如何將局部可交換性納入階層混合模型。最後，我們將簡單討論如何將主題模型應用於其他應用，並顯示分析部分帶註影像的結果。自始至終，我們著重於相關模型之圖形模型建構。圖形模型為主題建模假定提供有效的圖解，此為研究人員建立主題模型延伸和更多應用之槓桿。



【圖 1】來自 20 主題 LDA 模型的所有 20 個主題的熱門詞彙。後驗機率從《科學》的 17,000 篇文章估算而得（在 20 世紀 90 年代發表的所有文章）。每個字詞的大小正比於其在主題中出現的機率。LDA 後驗機率使用開放源碼軟體，採用 Gibbs 抽樣計算而得 [20]。

潛在狄氏配置

潛在狄氏配置 (LDA) 是階層機率模型，用於將文件集合分解為其多個明顯主題，LDA 的「主題」是一個詞彙的機率分佈 [15]。將 LDA 與其關聯事物稱為機率主題模型。

LDA 假設文件集中具有固定數量的主題，並假設每一篇文章都可視為這些主題的混合。在這些假定下分析文件集合時，機率推論演算法揭示一個內嵌主題結構（見圖 1）。透過此結構，LDA 提供了一種迅速摘錄摘要，探索和搜尋大規模文件集合的方式。

主題 β 是語彙 V 的某固定詞彙之分佈，Dirichlet 分佈是多項式參數向量之分佈，亦即其值為正值，且其和為 1 的向量。LDA 的生成機率假設假定文件集合的生成過程，如下所示（圖 2）：

- 1) 從對稱 Dirichlet 分佈推論 K 個主題， $\beta_k \sim Dir_V(\eta)$ ， $K \in \{1, \dots, K\}$ 。
- 2) 對於每一篇文章 d ，從對稱 Dirichlet $\theta_d \sim Dir_K(\alpha)$ ， $d \in \{1, \dots, D\}$ 推論主題比例。
- 3) 對於每一篇文章 d 中的每一個字詞 n ，
 - 基於主題所佔比例，推論主題分組， $z_{d,n} | \theta_d \sim Mult(\theta_d)$ 。
 - 從相應主題推斷字詞， $\omega_{d,n} | z_{d,n}, \beta_{1:K} \sim Mult(\beta_{z_{d,n}})$ 。

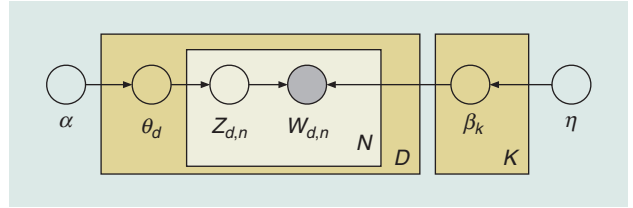
圖 2 中的圖形模型揭示 LDA 假設之套疊多層結構。LDA 由混合模型的階層建構。每一篇文章都透過有限混合模型進行建模，其中每篇文件的混合比例皆可唯一推斷，但混合元件（亦即主題）則為集合共用。就統計觀點而言，這被稱為歸屬度或混合成員模型 [25]。LDA 以心理學 [23] 和機器學習 [35] 的開創性工作為基礎，並與經典主成分分析方法聯繫密切 [18]。

生成過程定義潛在變數（主題、主題比例和主題分配）與觀測變數（字詞）之聯合機率分佈。我們透過檢定潛在變數已知觀測值之後驗分佈分析文件集合：

$$p(\beta_{1:K}, \theta_{1:D}, z_{1:D,1:N} | \omega_{1:D,1:N}) \quad (1)$$

主題 $\beta_{1:K}$ 後驗分佈之眾數識別字詞語料範圍之模式；主題比例之後驗分佈之眾數 θ_d 確定第 d 篇文章如何表達這些模式；主題分配之後驗分佈之眾數 $z_{d,n}$ 確定第 d 篇文章之第 n 字詞與哪一個主題關聯。

後驗推論可視為「逆」生成過程。視語料而定，後驗推論的目標是尋找生成其文件的替代主題結構之後驗分佈。此後驗機率提供無法在單獨字詞中獲得的架構，並且此架



【圖 2】LDA 之圖形模型表示 [15]。

構可以被用於探索和總結文件集合的架構。例如，圖 1 舉例說明來自 20 主題 LDA 後驗機率，這些機率由《科學》的文章推估而得。當與語言契合時，LDA 的後驗分佈與主題及關聯度之人類認知符合 [32] [52]，而且對文件和語料理解來說是可解釋的 [22] [44]。

精確計算式 (1) 的後驗機率很難操作 [15]。近似計算後驗機率是 LDA 之核心計算問題，構想和改進後驗機率近似計算演算法是主題建模研究之活躍領域。多數方法依靠變分推論 [15] [56] 或馬爾可夫鏈抽樣 [32]。在變分推論中，透過最佳化接近的分佈近似計算後驗機率。在馬爾可夫鏈抽樣方法中，定義馬爾可夫鏈之穩定分佈是 LDA 的後驗機率，長時間運轉馬爾可夫鏈，並且從平穩分佈（我們希望如此）收集樣本。

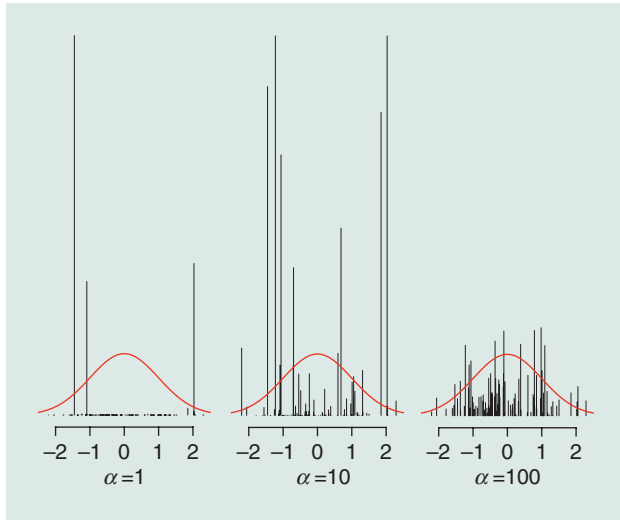
這些演算法慮及大規模文件集合分析，而且加速 LDA 推論成為機器學習研究的一個活躍領域。推論演算法的進展包括圖表處理器上的處理過程 [66]，機器間的平行處理 [43]，變化與抽樣法之組合 [63]。再者，LDA 推論的多個實作是免費的開放原始碼軟體 [8] [20] [38]。（圖 1 之模型符合 [20]。）

階層 DIRICHLET 過程

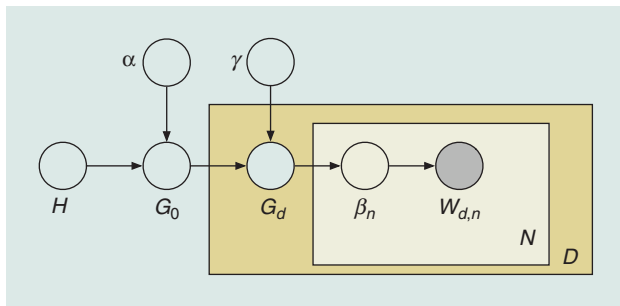
雖然我們可以使用 LDA 順利學會文件集合之主題，LDA 分析要求事先固定主題之數量（例如，圖 1 中的主題數量固定在 20）。這是個嚴重缺陷，透過潛在變數實現階層建模的原則之一是讓資料報知暗藏之結構。選擇 LDA 之主題數量通常藉由審查自行提出文件 [15] 之適合度完成，或根據整體集合 [32] 的邊際機率進行選擇 [36]。

作為將事前機率加於機率分佈無窮維空間之上的方法，貝氏非參數方法提供一種解答這個問題的嚴謹方法。特別是可在主題模型中使用離散 HDP，其中主題的數量事先是「無限大」[54]。當限定文件集合時，其展示的主題數量將成為潛在架構的之後驗分佈的一部分

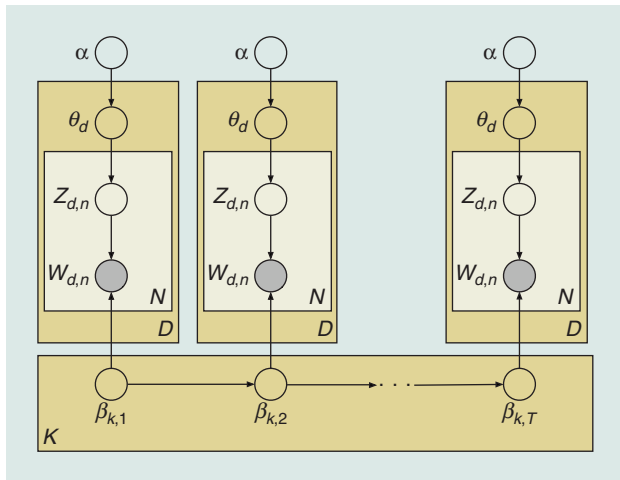
此外，HDP 慮及先前未見之文件會「導致」先前未見之主題。這對分析不斷生長與變化的集合特別有吸引力。即便是在有限樣本基礎之上選擇主題數量的優秀方法，也



【圖 3】具標準正態基本分佈之 DP 的三幅圖。DP 圖為離散型；隨其增加，產生的隨機分佈看起來更類似基本分佈。



【圖 4】HDP 主題模型之圖形模型表示 [54]。



【圖 5】DTM 之圖形模型表示 [12]。

經常毫無道理地假設所有未來之文件只使用那些主題。因為應用機率模型，HDP 允許新主題自然湧現。

在描述 HDP 主題模型之前，我們回顧 Dirichlet 過程 (DP) [5] [29]。DP 提供任意空間的多個分佈之上的分佈，可以表示為

$$G \sim DP(\alpha, G_0) \quad (2)$$

在此表達式中， G 本身就是某些空間上的分佈，其分佈有兩個參數：精確度參數 α 為正純量，以及與 G 在相同空間，分佈已知的基分佈 G_0 。（我們將在後續說明這些參數的作用。）DP 可用做各個分佈在無窮維空間上的先驗分佈。因此在這些設定中，稱之為貝氏非參數方法 [41]。

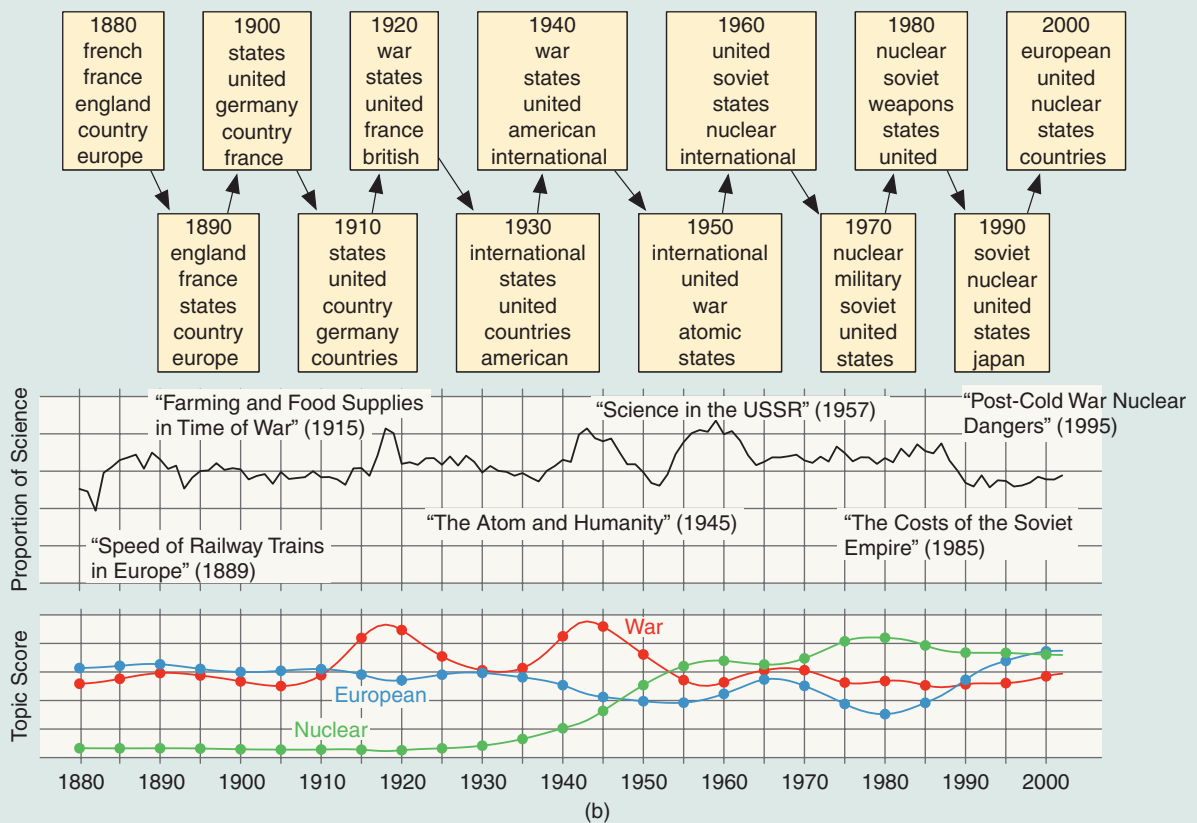
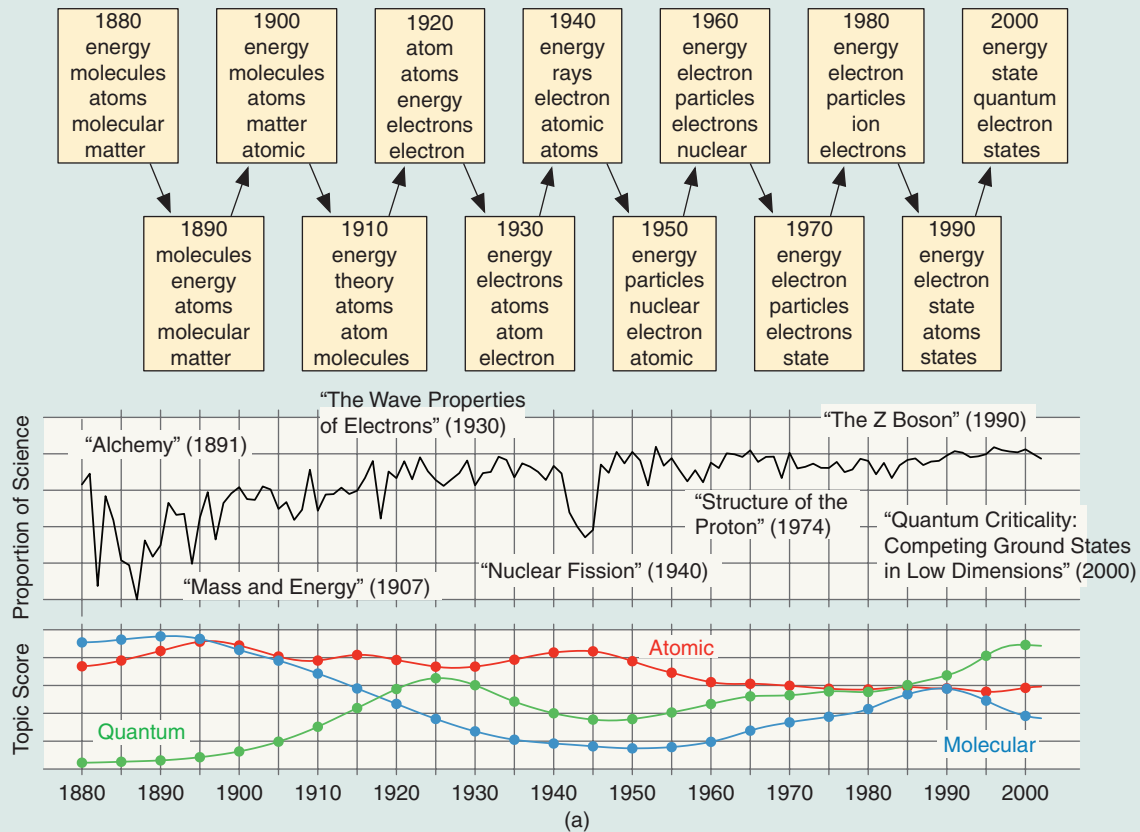
對於建構貝氏非參數主題模型而言，DP 有兩個重要的特性。首先，來自 DP 的抽取是離散型，在獨立產生自 G_0 的值（稱之為「原子」）處具有正機率質量函數。如果 G_0 是連續的，或包含無限多原子，則來自 DP 的每次抽取都為無限多個值分配非零機率值。精確度參數 α 控制整個原子的機率分配， α 較小意味著少數優勢原子擁有幾乎所有的機率。較大的 α 值則導致各個離散分佈更多與基分佈 G_0 類似。有關該特性的說明參見圖 3。

例如，假設 G_0 是具有零均值和單位變異數的標準常態分佈。隨機分佈 G 分佈在實數線上，在抽取自 G_0 的可數無限點集處有正機率質量函數。如果 α 較大，則這種分佈看起來像一個標準常態分佈的離散型，如果 α 較小，那麼個別點的機率較大。

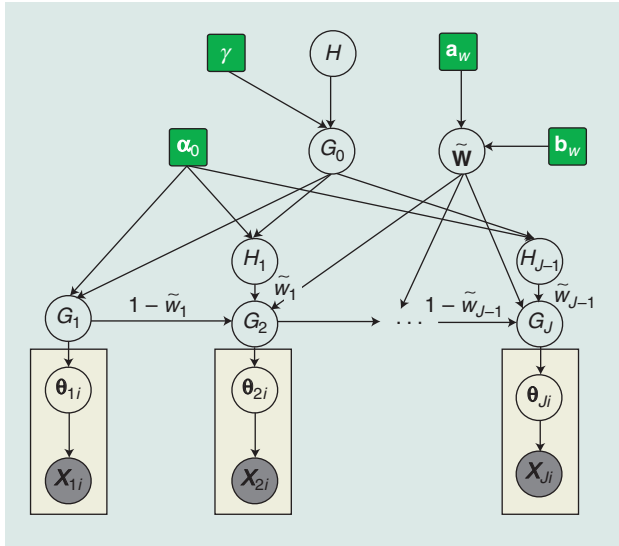
源自 DP 之抽取的第二個特性即所謂的「群集特性」。如果我們從 DP 抽取 G ，然後反覆多次從 G 中抽取變數，那麼所繪的圖將呈現一個分區結構，我們可以根據哪些分區共用相同原子而將他們分開。在上述範例中，這意味著從 G 反覆抽取 X_n 會形成一個實數集合，其中一些則完全相同的。請注意，從標準常態分佈中重複抽取則不具有此特性。

這一觀點闡明了該基分佈 G_0 以及尺度參數 α 的作用。所抽取的唯一值（即原子）與來自 G_0 的抽取相互獨立，如果 G_0 是連續的（例如在標準常態分佈範例中），則唯一值的個數隨著 α 增大而增大。雖然我們專注於這裡所述的隨機分佈表示，群集效應將 DP 與隨機劃分聯繫在一起，特別是 Chinese restaurant 過程 [45] 和 Ewens 抽樣公式 [27]。

現在我們回到主題建模。在 LDA 中，主題比例是 K 主題的分佈。為了建立一個非參數貝氏主題模型，我們將來自有限 Dirichlet 分佈的主題比例 θ_d 替換為來自 DP 的主題 G_d 的分佈。 G_d 的原子就是主題，即詞彙的多項式參數，因而基分佈 G_0 就是主題的分佈（在上面的範例中，這取



【圖 6】與《科學》檔案 (1880–2002) 符合的 DTM 的兩個範例主題。敬請注意，模型假定每個出版年份作一個時期。我們有說明每個十年的熱門字詞。重印這幅圖得到參考文獻 [13] 的許可。



【圖 7】dHDP 之圖形模型

代了高斯基分佈)。如要繪製一個文件，我們首先繪製分佈 $G_d \sim DP(\alpha, G_0)$ 。然後，對於每個字詞，我們從 G_d 抽取一個主題 β_n ，最後從 β_n 抽取字詞。群集特性至關重要，保證了文件的字詞將共用一個較小的主題子集。

然而，這並不足以充分指定貝氏非參數主題模型。LDA 的一個顯著特點是主題自身在語料間共用。如果基分佈 G_0 是主題的連續分佈，例如，對稱 Dirichlet 分佈，那麼文件中的字詞將共用相同主題，而文件間的字詞並非如此。在 HDP 中，基分佈 G_0 本身來自 DP， $G_0 \sim DP(\gamma, H)$ 。主題 G_d 的文件先驗分佈因而在文件間共用 [54]。

總而言之，HDP 主題模型透過下面的過程生成一個文件集合。

- 1) 決定主題的基分佈 $G_0 \sim DP(\gamma, H)$ ，其中 H 是字詞單純度的對稱 Dirichlet 分佈。
- 2) 對於每篇文件 d ，決定主題的文件先驗分佈 $G_d \sim DP(\alpha, G_0)$ 。
- 3) 對於每篇文件 d 中的每個字詞 n
 - 決定字詞的主題 $\beta_{d,n} \sim G_d$ 。
 - 決定字詞的分佈 $\omega_{d,n} \sim Mult(\beta_{d,n})$ 。

如圖 4 中圖形模型所示。

理論上講，HDP 主題模型之後驗分佈提供每個主題的文件分佈 G_d ，以及主題之語料先驗分佈 H 。然而在實際中，隨機分佈本身已邊緣化，後驗推斷提供主題索引（就 LDA 而言），以及字詞對應分佈的後驗分佈。關鍵區別是 HDP 中主題的數量由資料決定。此外，正如我們上面所提

到，在執行預測時，一個新文件可以「開啓」一個新主題，僅僅認為其是一個尚未出現的原子。

至於 LDA，對 HDP 的精確推斷非常棘手，已發展出若干方法。原有的近似推理演算法以馬可夫鏈蒙特卡羅法 (MCMC) 採樣 [54] 為基礎，但最近已提出速度更快的變分法 [55]。相較於 LDA，文獻 [54] 表明，使用基於邊緣相似度的模型選擇方法時，HDP 主題模型可以找到集合中「正確」的主題數目。

模組化循序群集

即便功能強大，LDA 假設文件可交換。也就是說，對於確定語料範圍內的主題，以及將每篇文件分解到這些主題時，其順序無關緊要。在本小節中，我們將描述兩個慮及循序文件群集的主題模型。這些方法模型語言以不同的方式隨時間變化。

動態主題建模

動態主題模型 (DTM) 慮及主題隨時間變化 [12]。假定語料按不同時段組織，每個時段都與自己的主題集合相關聯，每個時段中的每個主題都從先前時段中的同一主題隨機漂移而來。

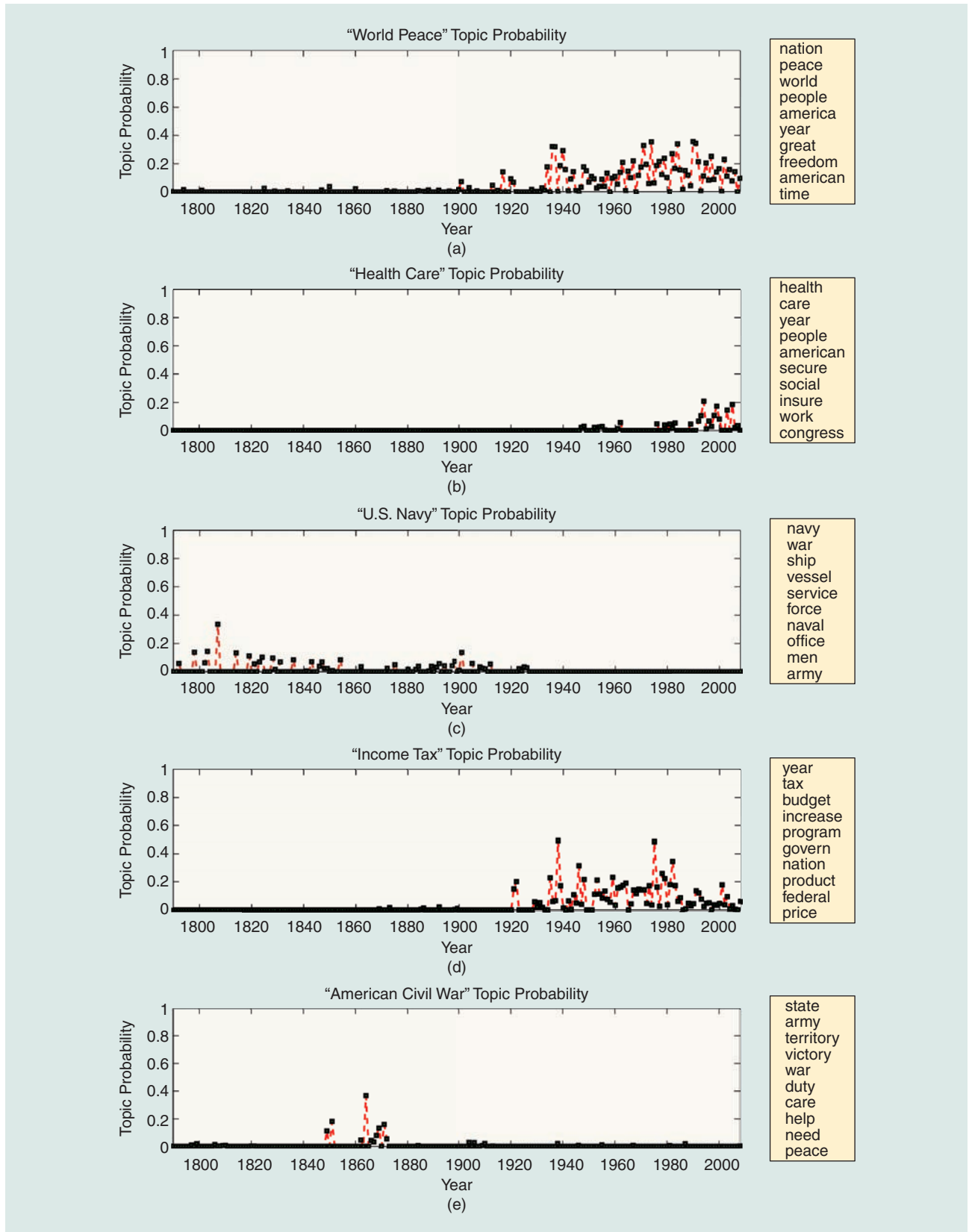
就特定時段的已知主題而言，此時段的文件的生成過程與 LDA 相同。讓我們回想一下，主題是字詞分佈的一個多項式參數。要建立 DTM，我們需要指定能夠反映每個主題如何在時段於時段間演變的一個模型。

多項式參數向量的循序模型之建置組是邏輯常態分佈。邏輯常態分佈是多項式參數的分佈，這是另一種較常用的 Dirichlet 分佈。參考文獻 [2] 首次慮及向量分量間相關性的複雜模式。其思想是從多元高斯中抽取一個實值向量，然後藉由先取指數再重歸一化轉化為一個多項式參數。

對於 DTM 而言，邏輯常態內嵌於一個狀態空間模型 [64]。主題單純度之共變異數矩陣是對角陣，即分量間相互獨立。然而，此分佈的均值為前一個時段實值隨機向量的均值。因此，要獲得時刻 t 的主題 K ，

$$y_{t,k} \sim N(y_{t-1,k}, I\sigma^2)$$

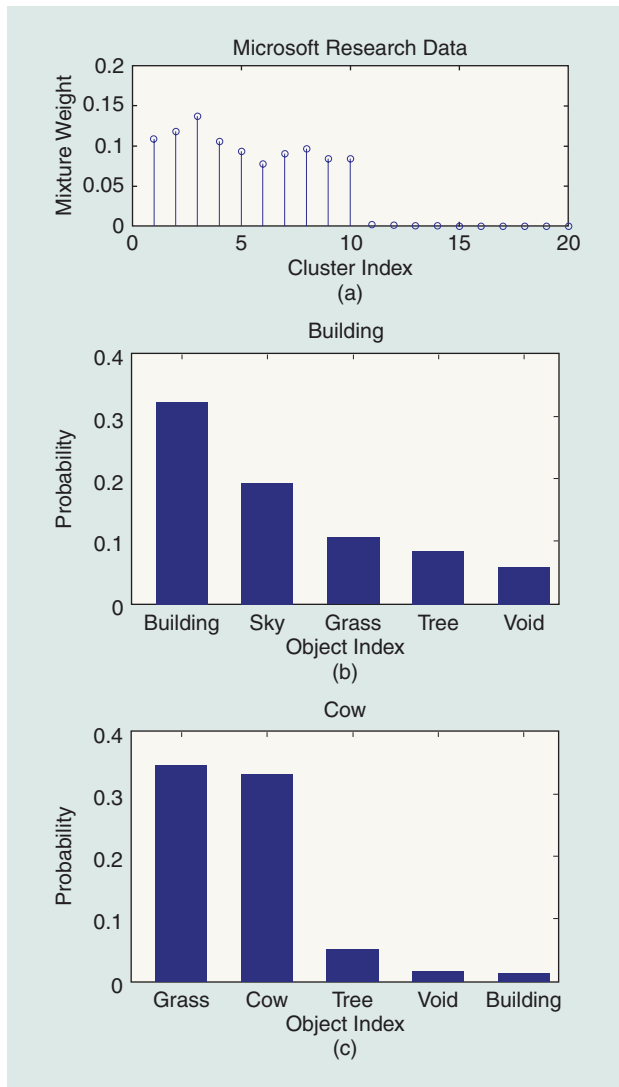
$$\beta_{t,k,v} = \frac{\exp\{y_{t,k,v}\}}{\sum_v \exp\{y_{t,k,v}\}} \quad (3)$$



[圖 8] 透過 dHDP 分析推斷得到的 1790 至 2008 年間美國總統國情諮文之主題。每幅圖將主題概率作為年份的函數進行描繪，而在每個主題中最有可能出現的字詞都顯示在右邊。概念主題：(一) 世界和平 (world peace)，(二) 醫療保健 (health care)，(三) 美國海軍 (U.S. Navy)，(d) 所得稅 (income taxes)，以及 (e) 美國內戰 (American Civil War)。此圖經許可轉載自參考文獻 [47]。

用 $\beta_{t,k,v}$ 表示在時刻 t 主題 K ，字詞 $v \in \{1, \dots, V\}$ 的概率。變異數參數 σ^2 控制每個字詞的對數機率之先驗漂移。請注意，DTM 和 LDA 的不同之處。在 LDA 中，每個字詞都與每個主題下機率相關聯。在 DTM 中，每個字詞是與每個主題下的一個機率序列相關聯。這些生成假設導向更豐富的后驗分佈。參見圖 5 的充分 DTM 模型之圖形模型表示。

透過逼近後驗分佈（參見 [12] 中的結構化變分法），更豐富的潛在空間允許以新方法來探討長期運行的循序語料庫。首先，我們可以透過查詢每個時段的主題之最熱門字詞，研究主題如何隨著時間而變化。其次，我們可以分析



【圖 9】推斷與 Microsoft 資料集相關聯之潛在特性。(a) 混合權重的後驗分佈，對概率場景主題進行量化（推斷十個主題）。(b) 及 (c) 部份顯示某限定類別的圖像或主題，物件機率的範例（物件/字詞的機率）；這裡我們只提出每個主題最熱門的五個字詞。此圖經許可轉載自參考文獻 [24]。

一個主題內特定詞的機率如何隨時間變化。最後，我們可以在不同的時間檢查與主題相關的文章。請注意，這就將文章相互關聯，同時考慮到字詞使用如何改變。傳統文件群集方法並不能解釋這樣的變化。

舉例來說，我們使用 DTM 分析 *Science* 雜誌（1880 至 2002 年）。這些文章由 JSTOR 服務掃描，並按年份分類。調查的兩個議題如圖 6 所示。

DTM 為群集提供 LDA 或 HDP 不曾有的窗口。另外 DTM 已經擴展到建模連續時間 [58]，以及作為文件影響模型的一個元件，即尋找高影響學術文獻的一種基於語言的方法 [31]。請注意，開發 DTM 的 HDP 類比是主題建模中一個開放問題。理想的情況下，隨著時間推移應當建模某種「生」與「死」的主題，也許還要分裂和合併主題。這種模型最初在參考文獻 [48] 中提出。

動態階層 DIRICHLET 過程

DTM 建模隨時間變化的主題。在本小節中，我們將描述基於時間之主題模型的不同類型，其中之一是建模主題比例隨時間變化。此模型，亦即動態 HDP (dHDP) [49]，是基於 HDP 的擴展 [54]，將在「階層 Dirichlet 過程」一節中對其進行討論。

如上所述，並非假設主題依賴於文件的分佈 G_a ，dHDP 假定主題分佈 G_t 對應時刻 t 所有的文件的分佈。我們的目標是強使 G_t 隨著時間的發展能夠大體平滑演進，具有及時應對急劇變化的潛力，可以根據需要描述資料。 G_t 依據如下假設得出：

$$G_0 = DP(\gamma, H) \quad (4)$$

$$\tilde{\omega}_t = Beta(a, b_t)$$

$$H_t \sim DP(\alpha_{0t}, G_0) \quad (5)$$

$$G_t = (1 - \tilde{\omega}_t)G_{t-1} + \tilde{\omega}_t H_t$$

其中 $\tilde{\omega}_1 = 1$ 。

如圖 7 所示之圖形模型。為了簡化問題，通常設定參數 (a, b_t) 是時間獨立，並設定具有高概率 $\tilde{\omega}_t$ 的參數比較小，但 $\tilde{\omega}_t$ 很少接近 1，這表明在某限定時刻，主題的機率有劇烈變化（例如，因為在某限定時刻，一個新的話題/主題突然變得很重要）。在 dHDP 中，由於 G_0 是離散的，每個 G_t 皆由相同的主題離散集合和在文件中使用每個主題隨時間演變之機率組成。透過此模型，我們再次從觀測資料以非參數方式推斷主題之數量，但已移除資料之時間互換之性質（與 HDP 相關）。

使用適當的近似，dHDP 模型利用變分貝氏推斷實施 [11]，並應用到 1790 至 2008 年間美國總統的國情諮文，此範例結果如圖 8 所示 [47]。主題之數量，成為重要主題之時段，以及主題相關的字詞機率均藉由從資料推斷而得，而人（作者）則定義與每個主題關聯（不完全）的標籤。

主題模型之其他應用

以上討論的範例主要集中於文件分析。然而主題模式已經推廣到許多其他類型的資料。比如，廣泛應用於電腦視覺問題、即分析圖像、視訊、圖像和文字之同步分析（例如標註）[3] [6] [10] [19] [24] [28] [37] [51] [53] [57]，[59]-[61] [65]。此外，研究人員已將主題模式應用於聲音特徵、基因標記[46]，調查回應 [26]，計算機碼 [4]，社交網絡資料 [1] [21] [42]。主題建模以發現文本結構的領域為起始，但已經成為在多種設定中探索群組資料的研究領域。

對於這些應用而言，圖形化模型的基本形成通常與相關聯的字詞/文件的模型非常相似，只是以不同的「字詞」形

式表現。例如，在圖像或視訊處理中，如果量化圖像之特徵，那麼碼簿每個成員（離散和有限）的形式就是這些字詞 [3] [24] [28] [51] [53]。在文件時間相關分析中，DP 中已移除可交換性假設，允許利用時間資訊（在時間上相近的文件可能被組成類似主題）。在圖像處理的應用中，圖片字詞（代碼）通常與圖像的局部區域相對應，而且已知圖片字詞間的相互關係的解釋又是可取的，現在透過其在圖像內的空間位置進行定義。具體來說，與圖像之相鄰區域相關聯的模型參數可能會更加密切相關。最近已經有一些文章著手解決此問題 [3] [24] [53]，明確地提出圖像範圍內特徵空間位置的解釋。

作為最近超越可互換性（「意象詞袋」）假設的一個範例，其中慮及羅吉斯 stick-breaking 過程 (LSBP) [24]。此 LSBP 與相關內核 stick-breaking 過程 (KSBP) [3]，使用一個空間內核來定義圖像中的親近度。徑向基函數是為此廣泛採用的內核。圖像特徵是否來自同一主題之機率透過依據空間核函數所構成之機率分佈定義。具體來說，在邏輯連



[圖 10] 範例分割和標記結果。(a) 原始圖像；(二) Corr-LDA [10]；(c) 使用 DP 進行圖片字詞分割；(四) 可見字詞的區塊在圖像 LSBP 中所佔空間位置 [24]。欄 1-3 來自微軟的資料集；欄 4-6 來自 UIUC 的體育資料集。所有這些範例並未在學習過程中進行標注。此圖經許可轉載自參考文獻 [24]。

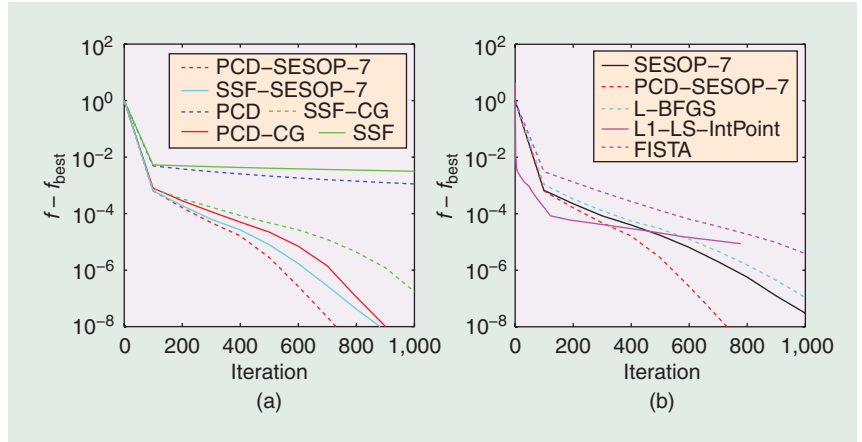
壓縮感測

第三個實驗是從圖像的隨機測量中恢復圖像，這也稱為壓縮感測。相對於先前的兩個實驗，我們對 64×64 像素的合成稀疏圖像 X 進行運算，只有 5% 的像素（隨機位置上）為非零值，出自獨立同分佈 (i.i.d.) 的高斯白色雜訊分佈 $\mathcal{N}(0, 10)$ 。這樣的圖像與天文或者雷達成像應用中所獲得的實際圖像相符。因為討論中的訊號是稀疏的，我們在公式中採用的字典是恒等運算元 $A = I$ 。

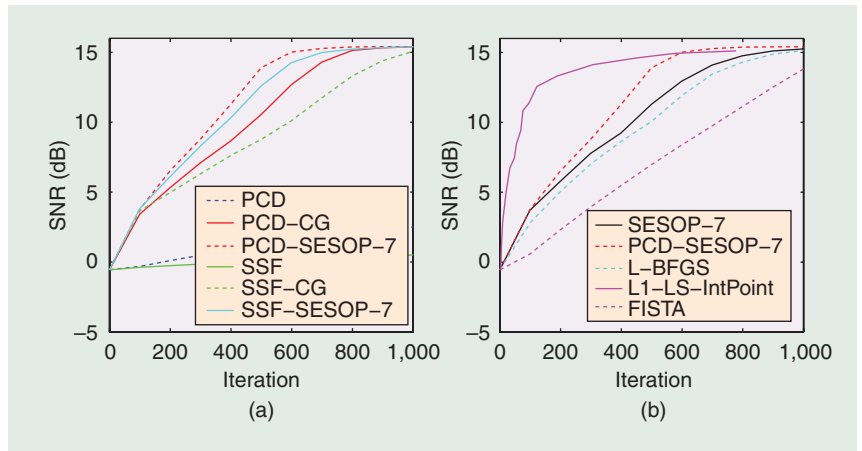
退化運算元 H (或 Q)，更符合「需要：稀疏和冗餘表示法」小節中的描述) 是由 410 個投影所構成，它們得自於隨機選擇對 X 套用二維快速傅葉爾轉換 (Fourier transform) 的 410 個樣本。因此， X 中的基數是 y 中測量值的兩倍，而我們的目標是恢復 x 。 y 中的值受到方差為 $\sigma = 0.01$ 的加性雜訊所污染。

圖 6 和圖 7 提出 $\lambda = 0.001$ 得出的結果，沿襲與先前實驗相同的結構。雖然大部分的演算法與之前的表現類似，L1-LS 卻表現出明顯差異，它在獲得高品質恢復結果方面快速有效。對比先前的兩個實驗，L1-LS 在此模擬中表現明顯不同的原因如下：在 z 中非零係數數目相對較少的情況下 (如本模擬)，預計 L1-LS 會十分有效，因為對數障礙函數的 Hessian 變成低秩矩陣，該秩對應於有效的約束條件數目 (訊號中的非零係數數目)。在本例中，用以解決牛頓系統的 inner-CG 方法變得十分有效。

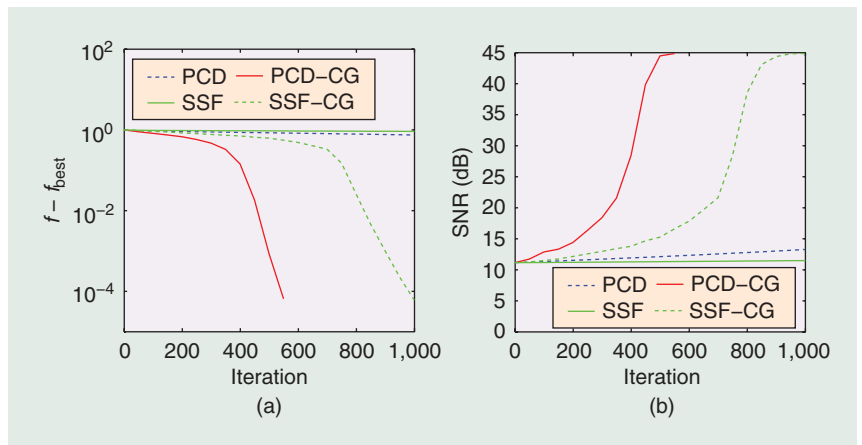
在目標函數中將 q_1 項替換為參數 $s = 1$ 的凹懲罰式(34)，可獲得大幅改善的結果。這樣會進一步促進結果的稀疏性，同更接近於理想圖像 X 。我們採用前一次用 PCD-SESOP-7 演算法進行 1000 次反覆運算的結果，並對同一組演算法進行初始化。圖 8 顯示精細化階段的目標函數和 SNR 值，透過 PCD-CG 演算法得到的結果具有明確且實質的改善。圖 9 顯示原始圖像、此



【圖 6】壓縮感測：對不同測試演算法的反覆運算次數函數與目標函數值。



【圖 7】壓縮感測：對不同測試演算法的反覆運算次數函數與 SNR 值。



【圖 8】壓縮感測：對於採用凹懲罰項的不同測試演算法，隨反覆運算次數函數變化的目標函數值和 SNR 值。

精細化階段的初始化與結果圖像，結果圖像看起來以近乎完美的方式重構指定圖像。

- [14] D. Blei and J. D. Lafferty, "Dynamic topic models," in *Proc. 23rd Int. Conf. Machine Learning*, 2006, pp. 113–120.
- [15] D. Blei, A. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.
- [16] A. Bosch, A. Zisserman, and X. Munoz, "Scene classification via pLSA," in *Proc. European Conf. Computer Vision*, 2006.
- [17] M. Braun and J. McAuliffe, "Variational inference for large-scale models of discrete choice," *J. Amer. Statist. Assoc.*, pp. 324–335, 2010.
- [18] W. Buntine and A. Jakulin, "Discrete component analysis," in *Subspace, Latent Structure and Feature Selection*. New York: Springer-Verlag, 2006.
- [19] L. Cao and L. Fei-Fei, "Spatially coherent latent topic model for concurrent segmentation and classification of objects and scenes," in *Proc. Int. Conf. Computer Vision*, 2007.
- [20] J. Chang. [Online]. Available: <http://cran.r-project.org/web/packages/lda/>
- [21] J. Chang and D. Blei, "Hierarchical relational models for document networks," *Ann. Appl. Statist.*, vol. 4, no. 1, 2010.
- [22] J. Chang, J. Boyd-Graber, S. Gerrish, C. Wang, and D. Blei, "Reading tea leaves: How humans interpret topic models," in *Advances in Neural Information Processing Systems*, vol. 22, Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, and A. Culotta, Eds. 2009, pp. 288–296.
- [23] S. Deerwester, S. Dumais, T. Landauer, G. Furnas, and R. Harshman, "Indexing by latent semantic analysis," *J. Amer. Soc. Inform. Sci.*, vol. 41, no. 6, pp. 391–407, 1990.
- [24] L. Du, L. Ren, D. Dunson, and L. Carin, "A Bayesian model for simultaneous image clustering, annotation and object segmentation," in *Proc. Neural and Information Processing Systems (NIPS'09)*, 2009.
- [25] E. Erosheva, "Bayesian estimation of the grade of membership model," in *Bayesian Statistics*. vol. 7, pp. 501–510, 2003.
- [26] E. Erosheva, S. Fienberg, and C. Joutard, "Describing disability through individual-level mixture models for multivariate binary data," *Ann. Appl. Statist.*, 2007.
- [27] W. Ewens, "The sampling theory of selective neutral alleles," *Theor. Popul. Biol.*, vol. 3, pp. 87–112, 1972.
- [28] L. Fei-Fei and P. Perona, "A Bayesian hierarchical model for learning natural scene categories," in *Proc. IEEE Computer Vision and Pattern Recognition*, 2005, pp. 524–531.
- [29] T. Ferguson, "A Bayesian analysis of some nonparametric problems," *Ann. Statist.*, vol. 1, no. 2, pp. 209–230, 1973.
- [30] A. Gelman, J. Carlin, H. Stern, and D. Rubin, *Bayesian Data Analysis*. London, U.K.: Chapman & Hall, 1995.
- [31] S. Gerrish and D. Blei, "A language-based approach to measuring scholarly impact," in *Proc. Int. Conf. Machine Learning*, 2010.
- [32] T. Griffiths and M. Steyvers, "Finding scientific topics," *Proc. Nat. Acad. Sci.*, 2004, vol. 101, pp. 5228–5235.
- [33] M. Hoffman, D. Blei, and P. Cook, "Content-based musical similarity computation using the hierarchical Dirichlet process," in *Proc. Int. Conf. Music Information Retrieval*, 2008.
- [34] M. Hoffman, D. Blei, and P. Cook, "Finding latent sources in recorded music with a shift-invariant HDP," in *Proc. Int. Conf. Digital Audio Effects*, 2009.
- [35] T. Hofmann, "Probabilistic latent semantic analysis," in *Proc. Uncertainty in Artificial Intelligence*, 1999.
- [36] R. Kass and A. Raftery, "Bayes factors," *J. Amer. Statist. Assoc.*, vol. 90, no. 430, pp. 773–795, 1995.
- [37] J. Li, R. Socher, and L. Fei-Fei, "Towards total scene understanding: Classification, annotation and segmentation in an automatic framework," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2009.
- [38] A. McCallum. [Online]. Available: <http://mallet.cs.umass.edu>
- [39] D. Mimno and A. McCallum, "Topic models conditioned on arbitrary features with Dirichlet-multinomial regression," in *Proc. Uncertainty in Artificial Intelligence*, 2008.
- [40] D. Mimno, "Reconstructing Pompeian households," in *Proc. Applications of Topic Models Workshop*, 2009.
- [41] P. Muller, S. Walker, N. Hjort, C. Holmes, Eds., *Bayesian Nonparametrics*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
- [42] R. Nallapati and W. Cohen, "Link-pLSA-LDA: A new unsupervised model for topics and influence of blogs," in *Proc. Int. Conf. on Weblogs and Social Media (ICWSM'08)*, 2008.
- [43] D. Newman, A. Asuncion, P. Smyth, and M. Welling, "Distributed inference for latent Dirichlet allocation," in *Proc. Neural Information Processing Systems (NIPS'07)*, 2007.
- [44] D. Newman, J. H. Lau, K. Grieser, and T. Baldwin, "Automatic evaluation of topic coherence," in *Proc. North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT'10)*, 2010.
- [45] J. Pitman, *Combinatorial Stochastic Processes* (Lect. Notes for St. Flour Summer School). New York, NY: Springer-Verlag, 2002.
- [46] J. Pritchard, M. Stephens, and P. Donnelly, "Inference of population structure using multilocus genotype data," *Genetics*, vol. 155, pp. 945–959, June 2000.
- [47] I. Pruteanu-Malinici, L. Ren, J. Paisley, E. Wang, and L. Carin, "Hierarchical Bayesian modeling of topics in time-stamped documents," *IEEE Trans. Pattern Anal. Machine Intell.*, 2009.
- [48] V. Rao and Y. W. Teh, "Spatial normalized gamma processes," in *Advances in Neural Information Processing Systems*, vol. 22, Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, and A. Culotta, Eds. 2009, pp. 1554–1562.
- [49] L. Ren, D. Dunson, S. Lindroth, and L. Carin, "Dynamic nonparametric Bayesian models for analysis of music," *J. Amer. Statist. Assoc.*, 2009.
- [50] C. Robert and G. Casella, *Monte Carlo Statistical Methods (Springer Texts in Statistics)*. New York, NY: Springer-Verlag, 2004.
- [51] J. Sivic, B. Russell, A. Efros, A. Zisserman, and W. Freeman, "Discovering object categories in image collections," *CSAIL*, MIT, Tech. Rep., 2005.
- [52] M. Steyvers and T. Griffiths, "Probabilistic topic models," in *Latent Semantic Analysis: A Road to Meaning*, T. Landauer, D. McNamara, S. Dennis, and W. Kintsch, Eds. Lawrence Erlbaum, 2006.
- [53] E. B. Sudderth and M. I. Jordan, "Shared segmentation of natural scenes using dependent Pitman-Yor processes," in *Proc. Neural Information Processing Systems (NIPS'08)*, 2008.
- [54] Y. Teh, M. Jordan, M. Beal, and D. Blei, "Hierarchical Dirichlet processes," *J. Acoust. Soc. Amer.*, vol. 101, no. 476, pp. 1566–1582, 2005.
- [55] Y. Teh, K. Kurihara, and M. Welling, "Collapsed variational inference for HDP," in *Proc. Neural Information Processing Systems (NIPS'07)*, 2007.
- [56] Y. Teh, D. Newman, and M. Welling, "A collapsed variational Bayesian inference algorithm for latent Dirichlet allocation," in *Proc. Neural Information Processing Systems (NIPS'06)*, 2006.
- [57] C. Wang, D. Blei, and L. Fei-Fei, "Simultaneous image classification and annotation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2009.
- [58] C. Wang, D. Blei, and D. Heckerman, "Continuous time dynamic topic models," in *Proc. Uncertainty in Artificial Intelligence (UAI)*, 2008.
- [59] X. Wang and E. Grimson, "Spatial latent Dirichlet allocation," in *Proc. Neural Information Processing Systems (NIPS'07)*, 2007.
- [60] X. Wang, X. Ma, and E. Grimson, "Unsupervised activity perception by hierarchical Bayesian models," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, June 2007, pp. 1–8.
- [61] X. Wang, X. Ma, and E. Grimson, "Unsupervised activity perception in crowded and complicated scenes using hierarchical Bayesian models," *IEEE Trans. Pattern Anal. Machine Intell.*, 2009.
- [62] X. Wang and A. McCallum, "Topics over time: A non-Markov continuous-time model of topical trends," in *Proc. 12th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, 2006, pp. 424–433.
- [63] M. Welling, Y. Teh, and B. Kappen, "Hybrid variational/Gibbs collapsed inference in topic models," in *Proc. Conf. Uncertainty in Artificial Intelligence (UAI'08)*, Citeseer, 2008, pp. 587–594.
- [64] M. West and J. Harrison, "Bayesian forecasting and dynamic models," in *Proc. Int. Conf. Machine Learning*. New York: Springer-Verlag, 1997.
- [65] O. Yakhnenko and V. Honavar, "Multi-modal hierarchical Dirichlet process model for predicting image annotation and image-object label correspondence," in *Proc. SIAM Conf. Data Mining (SDM'09)*, 2009.
- [66] F. Yan, N. Xu, and Y. Qi, "Parallel inference for latent Dirichlet allocation on graphics processing units," in *Advances in Neural Information Processing Systems*, vol. 22, Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, and A. Culotta, Eds. 2009, pp. 2134–2142.

訊號和影像處理中的

L1-L2 最佳化方法

反覆運算收縮及擴展

稀

疏和冗餘表示法為訊號提供強大的新興模型。該模型採用來自於預定的和過完備字典的一些原子的線性組合來逼近資料來源。通常這類模型適於求解資料的混合 ℓ_1 - ℓ_2 凸最佳化問題。反覆

運算收縮演算法建立一系列對於

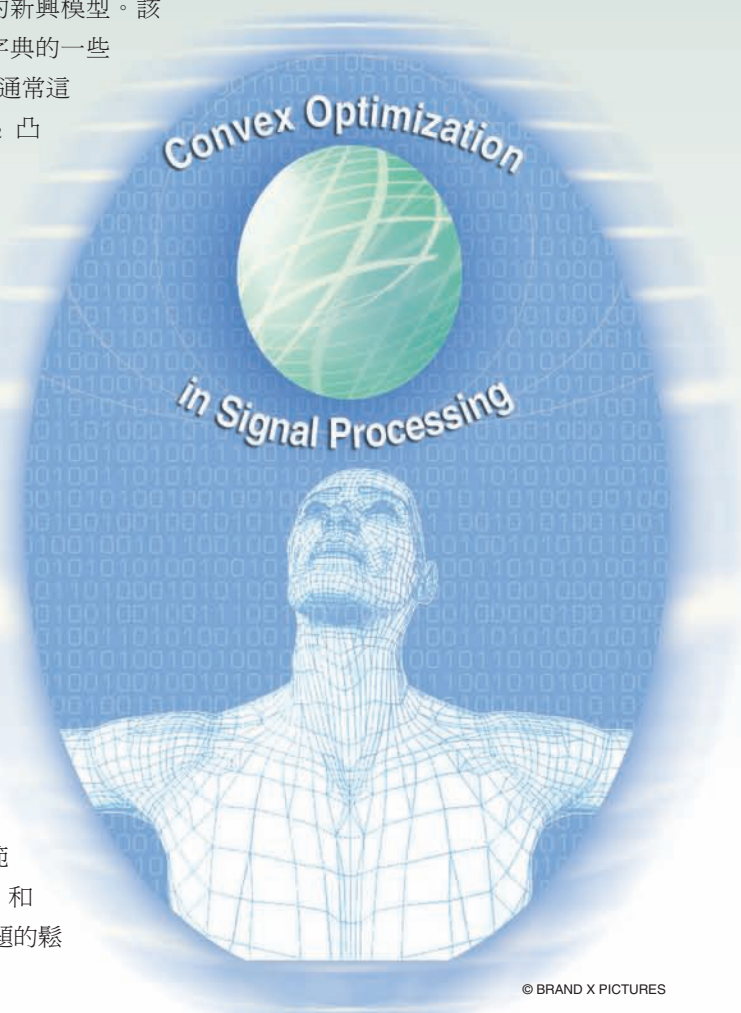
解決這些問題來說非常有效之數值方法，而且它們優於傳統的最佳化技術。在本文中，我們將概述這類方法，並從中衍生出基於順序子空間的最佳化演算法 (SESOP)、快速反覆運算軟閾值演算法 (FISTA) 和共軛梯度演算法 (CG) 的進展，呈現性能表現比較之餘也討論各種應用潛力，例如壓縮感測、層析成像和去模糊。

引言

在許多訊號建模的數學方法中，一種新穎且頗具吸引力的方法採用稀疏和冗餘表示 (參閱文獻 [1] 和 [2])。在這種模型中，訊號 $\mathbf{x} \in \mathbb{R}^n$ 被假定為來自於預定和冗餘字典 $\mathbf{A} \in \mathbb{R}^{n \times m}$ 的一些原子的線性組合，即，

$$\mathbf{x} = \mathbf{A}\mathbf{z} \quad (1)$$

這裡 $\mathbf{z} \in \mathbb{R}^m$ 假定是稀疏的， $\|\mathbf{z}\|_0 \ll n$ 。使用 ℓ_0 「範數」標記是為了表示非零向量 \mathbf{z} 的個數。給定一個訊號 \mathbf{x} 和字典 \mathbf{A} ，原子分解問題以尋找 \mathbf{x} 的稀疏表示為目標。該問題的鬆弛型允許較小的表示偏差，致使該問題變為：



© BRAND X PICTURES

$$\min_{\mathbf{z}} \|\mathbf{z}\|_0 \quad \text{subject to } \|\mathbf{x} - \mathbf{A}\mathbf{z}\|_2 \leq \epsilon \quad (2)$$

這裡 ϵ 是指表示原始訊號 \mathbf{x} 之 $\mathbf{A}\mathbf{z}$ 所允許的偏差。該問題是眾所周知的 NP-hard 問題 [1]，因此通常考慮使用逼近演算法。這類方法中較有吸引力的是基底搜尋 BP 演算法 [3] 或者 LASSO 演算法 [4]，將 l_0 替代為 l_1 範數。各種令人鼓舞之理論結果在於該問題的充分稀疏解唯一性，最稀疏解等價於最小 l_1 範數解，所有尋找訊號表示 \mathbf{z} 之建議是可行的計算任務 [1] [2] [5]-[8]。

在實用性方面，近年已有系列論文成功地將這些模型應用在不同的訊號和影像處理應用上，例如去噪、修補、去模糊、壓縮感測、源分離等等，這些應用都產生很先進的成果 [9]-[12], [14]-[16]。

許多應用都導向如下形式的 l_1 和 l_2 混合運算式最佳化問題

$$f(\mathbf{z}) = \frac{1}{2} \|\mathbf{x} - \mathbf{A}\mathbf{z}\|_2^2 + \lambda \|\mathbf{z}\|_1 \quad (3)$$

這裡 f 為向量 \mathbf{z} 的函數， \mathbf{x} 和 \mathbf{A} 假定是已知的。該問題是公式 (2) 的變形之一，這裡的約束條件由懲罰因數代替。參數向量 λ 取代 (2) 中的閾值 ϵ ，用於控制表示誤差和稀疏性之間的折中。我們應該注意，雖然這裡用 l_1 範數測量 \mathbf{z} 的稀疏性，但其實也可以使用很多類似的附加測量 $\rho(\mathbf{z})$ ，因此本文中大部分的公式推導都保持相同 (更多細節請參閱文獻 [17])。在不同的應用中，目標都是得到 $f(\mathbf{z})$ 對 \mathbf{z} 的最小化。因此在這些應用中，都是靠求解凸規劃來達成稀疏和冗餘表示。

直到最近，這種形式的最佳化問題都使用以下傳統方法進行處理：1) 各種經典的反覆運算最佳化演算法，範圍從最速下降法 (CG) 和反覆運算重加權最小平方方法，到所有涉及內點演算法的方法 [3]；2) 同倫法 [18], [19]；3) 貪婪演算法 [20]。然而這些方法通常效率不高，需要許多次反覆運算和過多的 CPU 時間才能達成求解。此外，同倫和貪婪演算法對高維問題 (例如 $m = 10^6$) 不適用，而這正是影像處理中經常會遇到的問題。

近年來，一類替代性數值演算法逐漸成形，它們在解決上述最佳化問題方面非常有效。這些演算法最初是作為一種啟發方法，用以解決有效運算方式方面的稀疏表示問題 [22]-[24]。這類方法後續跟隨有更多的推導和分析 [25]-[33], [17]。作為反覆運算收縮演算法，它們是經典 Donoho-Johnstone 收縮方法的擴展 [34]-[36]。大致來說，在這些方法中，每個反覆運算包含了 \mathbf{A} 及其伴隨矩陣的乘積，隨著尺度收縮步驟來獲得 \mathbf{z} 。儘管它們構造簡單，但在 (3) 中 $f(\mathbf{z})$ 的最小化方面，事實證明這些方法非常有效。充分的理論分析可證明這些方法的收斂性，可確保

此解是凸函數 f 之全域最小值，還對這些演算法展現的收斂速率進行研究，並且提出有助進一步改善的快速演算法 [26] [37] [17]。

需要：稀疏和冗餘表示

如果我們接受稀疏和冗餘表示模型的基本假定，包括來自於乘式 $\mathbf{x} = \mathbf{A}\mathbf{z}$ 的訊號 \mathbf{x} 與非常稀疏的 \mathbf{z} 。如何才算是「非常」稀疏？這個問題沒有明確的答案，但是我們可以假設如下：1) \mathbf{z} 是線性等式 $\mathbf{x} = \mathbf{A}\mathbf{z}$ 的最稀疏解，2) 它可以通過實用演算法重新構造，如 BP [5][6]，亦即 BP 演算法的解可精確恢復 \mathbf{z} 。下麵我們針對不同應用，給出基於上述假設的模型建構方法。

去噪

慮及含有雜訊形式 \mathbf{x} ，是由零均值獨立同分佈 (i.i.d.) 的高斯向量 $\mathbf{v} \in \mathbb{R}^n$ 得到，也就是說，我們觀測到 $\mathbf{y} = \mathbf{x} + \mathbf{v}$ ，再從中恢復 \mathbf{x} 。假定向量 \mathbf{z} 符合獨立的 Laplacian 分佈 $P(\mathbf{z}) \sim \exp\{-\lambda \|\mathbf{z}\|_1\}$ ，最大事後機率 (MAP) [3] 估計可將此去噪任務導向最佳化問題，

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{z}\|_2^2 + \lambda \|\mathbf{z}\|_1 \quad (4)$$

由 $\hat{\mathbf{x}} = \mathbf{A}\hat{\mathbf{z}}$ 可得到結果。如我們所見，我們面對的目標函數是 (3) 式中的 $f(\mathbf{z})$ 。

廣義 (線性) 逆問題

類似於上述內容，假定原始訊號 \mathbf{x} 經過線性退化操作 \mathbf{H} ，然後添加了上述的雜訊。這意味著我們觀測到 $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$ ，目標是恢復 \mathbf{x} 。MAP 估計會獲得如下運算式：

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z}} \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{A}\mathbf{z}\|_2^2 + \lambda \|\mathbf{z}\|_1 \quad (5)$$

估計結果還是之前的 $\hat{\mathbf{x}} = \mathbf{A}\hat{\mathbf{z}}$ 。這時我們獲得式 (3) 準確的目標函數和有效字典 $\mathbf{H}\mathbf{A}$ 。該構造可以為不同的任務服務，例如去噪 (這時 $\mathbf{H} = \mathbf{I}$)，去模糊 (這裡 \mathbf{H} 是模糊操作)，X 線斷層攝影中的反 Randon 變換 (這裡 \mathbf{H} 是投影運算元)，以及訊號插值 (去馬賽克、圖像修補)，這裡 \mathbf{H} 表示不同的抽樣/模糊化模板或其它。

訊號分離

假定我們指定一個混合訊號 $\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3 + \mathbf{v}$ ，這裡 \mathbf{v} 是加性雜訊， \mathbf{x}_1 來自於所採用字典 \mathbf{A}_1 的模型， \mathbf{x}_2 與使用字典 \mathbf{A}_2 模型有關。我們希望將訊號分離成三個分量。再次利用 MAP 得到 [12]：

$$\hat{z}_1, \hat{z}_2 = \arg \min_{z_1, z_2} \frac{1}{2} \|y - A_1 z_1 - A_2 z_2\|^2 + \lambda (\|z_1\| + \|z_2\|) \quad (6)$$

由 $\hat{x}_1 = A_1 \hat{z}_1$ 和 $\hat{x}_2 = A_2 \hat{z}_2$ 得到這些估計。上述問題，如形態成分分析 (MCA) [12]，再次符合式 (3) 所描述的廣義結構，只需用 $z = [z_1^T, z_2^T]^T$ 和 $[A_1, A_2]$ 取代 (3) 式中的 A 。

壓縮

指定一個訊號 x ，我們的目標是將其壓縮，並允許一些預先指定的 ℓ_2 範數誤差 ϵ 。這表示我們將它的運算式 \hat{z} 做如下求解：

$$\hat{z} = \arg \min_z \|z\| \quad \text{subject to } \|x - Az\|_2^2 \leq \epsilon^2 \quad (7)$$

當尋找關於近似不等式中滿足上述約束條件的 λ 時，求解該問題可以用求解式 (3) 中描述的問題來代替。實際壓縮的完成是靠量化向量 \hat{z} 中的非零項，然後沿著標誌位元發送其值。我們發現自己再次使用式 (3) 中的解決方式。

壓縮感測

假定我們以壓縮時訊號 x 的抽樣為目標。壓縮感測建議我們感測向量 $y = Qx$ ，這裡 $Q \in \mathbb{R}^{q \times n}$ 包含一系列 $q \ll n$ 的投影方向，訊號分量可以看作是向這些方向上投影 [14][15]。訊號從樣本中重構是透過求解如下算式所得：

$$\hat{z} = \arg \min_z \frac{1}{2} \|y - QAz\|^2 + \lambda \|z\|$$

而且 $\tilde{x} = A\hat{z}$ (8)

我們再次需要求解式 (3)。

以上範例表明式 (3) 描述的結構是所有這些 (與其他) 應用的根本基礎，因此我們需要利用高效率演算法來求解這一最佳化問題。

歸一化案例—靈感來源

歸一化案例的封閉式解

假如 A 是歸一化的，則可以透過系列簡單步驟操作 $f(z)$ 的最小化，轉化成一組 m 個獨立且相同的一維 (1-D) 最佳化問題，這比較容易解決。從公式 (3) 出發，採用恒等式 $AA^T = I$ ，我們可以得到：

$$\begin{aligned} f(z) &= \frac{1}{2} \|x - Az\|_2^2 + \lambda \|z\| \\ &= \frac{1}{2} \|A(A^T x - z)\|_2^2 + \lambda \|z\| \end{aligned} \quad (9)$$

利用 ℓ_2 範數為酉不變的事實，我們可以移除第一項中的乘數 A 。記 $z_0 = A^T x$ ，可得：

$$\begin{aligned} f(z) &= \frac{1}{2} \|z - z_0\|_2^2 + \lambda \|z\| \\ &= \sum_{k=1}^m \left[\frac{1}{2} (z_0[k] - z[k])^2 + \lambda |z[k]| \right] \end{aligned} \quad (10)$$

可以使用下式輕易得出關於 τ 之純量函數 $g(\tau) = 0.5(\tau_0 - \tau)^2$ 的最小化：

$$\tau_{\text{opt}} = \begin{cases} 0 & |\tau_0| \leq \lambda \\ \tau_0 - \text{sign}(\tau_0)\lambda & \text{otherwise} \end{cases} = \mathcal{S}_\lambda(\tau_0) \quad (11)$$

此函數將接近原點的值映射至零附近，其它值亦向零點收縮，這即便是運算子名稱的來歷。

回到我們最初的問題上，我們根據如下兩個步驟找到了最小化的封閉式解：1) 計算 $z_0 = A^T x$ ，2) 將 \mathcal{S}_λ 代入 z_0 得到 \hat{z} 。原本的問題提升為：當酉矩陣轉換為非酉 (或者非方陣) 矩陣時，我們是不是一定會失去這種簡單性？要回答這個問題，我們首先要考慮在某種程度上更廣義的形式的 A ——酉矩陣的串接，其中透過反覆執行收縮步驟實現。

區塊座標鬆弛演算法

當把 A 構建成多酉矩陣的聯集時發生什麼狀況？Sardy 等人在文獻 [38] 中提出答案，結果導向他們的區塊座標鬆弛演算法 (BCR)。我們在這裡簡單介紹該方法，將它看做是上述歸一化案例求解的直接擴展。為了簡單起見，我們可以假設 $A = [\Psi, \Phi]$ ，其中 Ψ 和 Φ 是 $n \times n$ 歸一化矩陣。輸出的最小化可以改寫為關於兩個長度為 n 的 z 部分，記為 z_Ψ 和 z_Φ ，

$$\begin{aligned} f(z) &= f(z_\Psi, z_\Phi) \\ &= \frac{1}{2} \|x - Az\|_2^2 + \lambda \|z\| \\ &= \frac{1}{2} \|b - \Psi z_\Psi - \Phi z_\Phi\|_2^2 + \lambda \|z_\Psi\| + \lambda \|z_\Phi\| \end{aligned} \quad (12)$$

BCR 演算法的核心觀念是將 $f(z_\Psi, z_\Phi)$ 關於 z 的兩個部分分別並交替最佳化，這正是 BCR 名稱的由來。假設我們保留第 k 次反覆運算時的當前解，記為 z^k ，由兩個部分構成： z_Ψ^k 和 z_Φ^k 。如果 z_Φ^k 保持不變，函數 $f(z_\Psi, z_\Phi)$ 可以寫為：

$$f(z_\Psi, z_\Phi^k) = \frac{1}{2} \|\tilde{x} - \Psi z_\Psi\|^2 + \lambda \|z_\Psi\| \quad (13)$$

其中 $\tilde{x} = x - \Phi z_\Phi^k$ 。這是與歸一化案例相同的函數，因此我們可以輕易得出封閉式解。此求解如下：

$$z_\Psi^{k+1} = \mathcal{S}_\lambda(\Psi^T \tilde{x}) = \mathcal{S}_\lambda(\Psi^T(x - \Phi z_\Phi^k)) \quad (14)$$

同樣地，一旦算出 x_Ψ^{k+1} 並保持不變，即可使用封閉運算式將相對於 x_Φ 的函數 $f(x_\Psi^{k+1}, x_\Phi)$ 最小化，方式如下：

$$x_\Phi^{k+1} = \mathcal{S}_\lambda(\Phi^T(b - \Psi x_\Psi^{k+1})) \quad (15)$$

採用這種方式，兩個更新階段之間的交替，此全域函數單調遞減，且可證明收斂於懲罰函數的全域最小值。

上述順序求解過程也可以採用替代方式，亦即平行執行兩個更新。如前所述，我們以第 k 次反覆運算的當前解著手 (z^k 由 z_Ψ^k 和 z_Φ^k 兩個部分構成)，使用以下方式提出兩個部分的平行更新：

$$\begin{aligned} z_\Psi^{k+1} &= \mathcal{S}_\lambda(\Psi^T(x - \Phi z_\Phi^{k+1})) \\ z_\Phi^{k+1} &= \mathcal{S}_\lambda(\Phi^T(x - \Psi z_\Psi^{k+1})) \end{aligned}$$

與前一個演算法唯一不同的是，在 x_Φ^{k+1} 的更新中，採用 x_Ψ^k 替代 x_Ψ^{k+1} 。要將這兩者併入一個公式，可以遵照下列簡單步驟，並利用收縮運算元在純量上獨立操作的事實。

$$\begin{aligned} z^{k+1} &= \mathcal{S}_\lambda \left\{ \begin{bmatrix} \Psi^T(x - Az^k + \Psi z_\Psi^k) \\ \Phi^T(x - Az^k + \Phi z_\Phi^k) \end{bmatrix} \right\} \\ &= \mathcal{S}_\lambda(A^T(x - Az^k) + z^k) \end{aligned} \quad (16)$$

爲了從一個反覆運算到下一個反覆運算來更新全域解，我們透過這種方式獲得感興趣的封閉解。當然，我們想要知道此公式是否可以用於更廣義的矩陣 A (及其變化)。令人驚喜的是，答案是肯定的，我們將在後續說明。

反覆運算收縮演算法

目前許多不同的反覆運算收縮演算法，它們之間有很重大的差異。這些方法的最初跡象出現在 Jansson 用於光譜和色譜測量系統的去卷積法中[21]，稍後由 Hoch 等人承繼這方面的工作，比如在 1990 年應用於核磁共振光譜學中的上下文最大熵估計 [22]。Starck 等人在 1995 年，也使用該演算法來考量普通圖像去模糊問題 [23]。Kingsbury 和 Reeves 在 2002 年接棒，他們採用類似方法處理複小波 [24]。然而，所有這些嘗試都缺乏與目標函數 (3) 之間的明確聯繫。

衍生反覆運算收縮演算法的出發點可能大爲不同，例如統計估計原理中的最大期望 (EM) 演算法 [25]、近似點和替代函數 [26],[27],[30]、定點策略的採用 [29]、平行座標下降

在許多對訊號進行數學建模的方式中，近來最具吸引力的方法是採用稀疏和冗餘表示。

演算法的運用 [28],[17]、貪婪演算法的變形 [39]、Bregman 反覆運算正則化 [32] 等等。這顯示此類演算

法多才多藝，可供我們爲上述最佳化任務設計求解器。

在文本中，我們將詳細介紹其中兩種演算法。我們刻意省略這些演算法的理論研究。然而，我們應該說明下列重要事實：這些演算法保證式 (3) 中的函數可收斂至全域最小值[26][17][31]。

此研究領域十分活躍，創新的反覆運算收縮方法幾乎日新又新，研究者使用它們進行分析或應用。我們在此僅提供部分研究清單[31]–[33], [40]，因爲我們沒有能力針對此「全速前進的列車」提供發展全貌和最新進展。

替代函數和鄰近點方法

我們以 Daubechies 等人在其建構 [26] 中使用的近似 (或替代) 目標開始思考 (3) 中的原函數，

$$f(z) = \frac{1}{2} \|x - Az\|_2^2 + \lambda \|z\|$$

我們將它加入下項：

$$\text{dist}(z, z_0) = \frac{c}{2} \|z - z_0\|_2^2 - \frac{1}{2} \|Az - Az_0\|_2^2$$

參數 c 的選擇用以滿足函數分佈爲嚴格凸(關於 z)，這意味著要求它的 Hessian 矩陣是正定的： $cI - A^T A > 0$ 。這由選擇 $c > \|A^T A\|_2 = \lambda_{\max}(A^T A)$ 來滿足 (矩陣 $A^T A$ 的最大特徵值)。新的目標函數變爲：

$$\begin{aligned} \tilde{f}(z, z_0) &= \frac{1}{2} \|x - Az\|_2^2 + \lambda \|z\| \\ &\quad + \|z - z_0\|_2^2 - \frac{1}{2} \|Az - Az_0\|_2^2 \end{aligned} \quad (17)$$

這是提出演算法採用的替代目標函數。如下面所示，項 $\|Az\|_2^2$ 從 \tilde{f} 中消失，這使得最小化問題變得更簡單。重新整理式 (17)，我們得出新的運算式：

$$\begin{aligned} \tilde{f}(z, z_0) &= \text{Const} - z^T [A^T(x - Az_0) + cz_0] \\ &\quad + \lambda \|z\|_1 + \frac{c}{2} \|z\|_2^2 \end{aligned} \quad (18)$$

在上述運算式中以我們稍後推導中的常量，其所包含的所有項都分別與 x 和 z_0 相關。定義下項：

$$v_0 = \frac{1}{c} A^T(x - Az_0) + z_0 \quad (19)$$

則上述的替代目標函數可以改寫為：

$$\begin{aligned}\tilde{f}(\mathbf{z}, \mathbf{z}_0) &= \text{Const} - c\mathbf{z}^T\mathbf{v}_0 + \lambda\|\mathbf{z}\|_1 + \frac{c}{2}\|\mathbf{z}\|_2^2 \\ &= \text{Const} + \lambda\|\mathbf{z}\|_1 + \frac{c}{2}\|\mathbf{z} - \mathbf{v}_0\|_2^2\end{aligned}\quad (20)$$

請注意，此公式和在式 (10) 中歸一化案例所得到的公式很類似。這意味著替代目標函數的最小值如下：

$$\mathbf{z}_{\text{opt}} = \mathcal{S}_{\lambda/c}(\mathbf{v}_0) = \mathcal{S}_{\lambda/c}\left(\frac{1}{c}\mathbf{A}^T(\mathbf{x} - \mathbf{A}\mathbf{z}_0) + \mathbf{z}_0\right)\quad (21)$$

至此，我們將原函數 f 轉化為新函數 \tilde{f} ，這樣我們就能得到全域最小值的封閉運算式。目標函數的變化取決於向量 \mathbf{z}_0 的選擇。採用替代目標函數的核心觀念是，我們針對函數 f 進行反覆地最小化運算，最後產生結果序列 $\{\mathbf{z}_i\}$ ，其中在第 $i+1$ 次反覆運算時將 \tilde{f} 最小化並指定 $\mathbf{z}_0 = \mathbf{z}_i$ 。值得注意的是，求解序列 $\{\mathbf{z}_i\}$ 會收斂至原函數 f 的最小值。因此，可輕易得出以下演算法：

$$\mathbf{z}_{i+1} = \mathcal{S}_{\lambda/c}\left(\frac{1}{c}\mathbf{A}^T(\mathbf{x} - \mathbf{A}\mathbf{z}_i) + \mathbf{z}_i\right)\quad (22)$$

我們此後提及此演算法時會稱為可分離的替代函數化方法 (SSF)。

上述方法可以理解為近似點演算法 [41]，這是最佳化理論中相當知名的方法 [17]。函數 $\text{dist}(\mathbf{z}, \mathbf{z}_0)$ 是前一個求解的距離測量。在第 i 次反覆運算時新增至原函數，可以增進反覆運算過程之序列估計彼此間的近似性。出人意料的是，原本預期會出現減緩效應，但在此處討論的案例中，實際上卻促成與傳統梯度下降法相對的加速效應。

上述討論可以用一套不同的研究取代，其中有一個導出很相似的演算法。此替代觀點出現在 Figueiredo 等人的研究工作中，他們使用最大期望 (EM) 估計或更好的方法，即替代確定性最佳化基礎——邊界最佳化方法 [25], [27], [30]。此方法也稱為極小化最佳化方法。

式 (3) 中函數 $f(\mathbf{z})$ 的最小化十分困難，因此邊界最佳化方法主張使用具備以下性質的相關函數 $Q(\mathbf{z}, \mathbf{z}_0)$ ：

- 1) 等式在 $\mathbf{z} = \mathbf{z}_0$ 時， $Q(\mathbf{z}, \mathbf{z}_0) = f(\mathbf{z}_0)$
- 2) 對所有的 \mathbf{z} ，原函數上限為 $Q(\mathbf{z}, \mathbf{z}_0) \geq f(\mathbf{z}_0)$

在特定的正則性條件下，由遞推公式產生求解序列：

$$\mathbf{z}_{i+1} = \arg \min_{\mathbf{z}} Q(\mathbf{z}, \mathbf{z}_i)\quad (23)$$

收斂於原函數 $f(\mathbf{z})$ 的一個局部最小值 [42], [43], [17]。由於這易於證明，來自於先前章節的項， $Q(\mathbf{z}, \mathbf{z}_0) = f(\mathbf{z}) + \text{dist}(\mathbf{z}, \mathbf{z}_0)$ 滿足上述兩個條件，同樣地導出所需的最佳化演算法。請注意，由於具備可分性， $Q(\bullet)$ 很容易最佳化。

平行座標下降演算法

我們現在說明平行座標下降 (PCD) 演算法 [28], [17]。它始於簡單的座標下降演算法，接著將這類下降步驟合併為一個更為簡單的聯合步驟，然後導出 PCD 反覆運算收縮方法。這個操作原理讓人想起前面提到的 BCR 演算法，以及我們為其所提及的平行更新修改。

回到式 (3) 中 $f(\mathbf{z})$ ，我們提出座標下降 (CD) 算法，其中每次只能更新一個 \mathbf{z} 輸入。這一系列 m 個步驟的迴圈 (訪問每個 $\mathbf{z} \in \mathbb{R}^m$ 的座標) 是必然收斂的。有趣的是，正如我們即將展示的，每個步驟都可以透過收縮得到，與先前描述的過程類似。

假設當前解是 \mathbf{z}_0 ，我們想要更新它當前值 $\mathbf{z}_0(k)$ 附近的第 k 個輸入。這會導出以下的一維函數：

$$g(\gamma) = \frac{1}{2}\|\mathbf{x} - \mathbf{A}\mathbf{z}_0 - \mathbf{a}_k(\gamma - \mathbf{z}_0[k])\|_2^2 + \lambda|\gamma|\quad (24)$$

向量 \mathbf{a}_k 是 \mathbf{A} 的第 k 欄向量。其中項 $\mathbf{a}_k(\gamma - \mathbf{z}_0[k])$ 會消除舊值的影響並加入新的值。經過幾次簡單的代數步驟會導出下列等式。

$$g(\gamma) = \frac{1}{2}\|\mathbf{a}_k\|_2^2 \cdot (\gamma - \gamma_0)^2 + \lambda|\gamma|\quad (25)$$

其中我們定義

$$\gamma_0 = \frac{\mathbf{a}_k^T(\mathbf{x} - \mathbf{A}\mathbf{z}_0)}{\|\mathbf{a}_k\|_2^2} + \mathbf{z}_0[k]$$

如同之前，我們找到自己的公式，它個類似於式 (10) 的公式，這意味著 $\mathbf{z}[k]$ 的最佳值為

$$\mathbf{z}_k^{\text{opt}} = \mathcal{S}_{\lambda/\|\mathbf{a}_k\|_2^2}\left(\frac{\mathbf{a}_k^T(\mathbf{x} - \mathbf{A}\mathbf{z}_0)}{\|\mathbf{a}_k\|_2^2} + \mathbf{z}_0[k]\right)\quad (26)$$

然而，採用該更新規則反覆更新 \mathbf{z} 的輸入，這對低維案例很有效，但對一些重要的應用卻不可行，例如對付大規模問題 ($> 10^6$)，採用矩陣 \mathbf{A} 的隱式表示。在這些應用中， \mathbf{A} 無法明確獲取，具有所謂可用函數的替代者，可快速地將 \mathbf{A} 或者 \mathbf{A}^T 應用於適合的向量。等式 (27) 需要一經要求就明確取得 \mathbf{A} 中的特定欄，因此不適合在隱式運算元設置中使用。

因此，我們想到上述方法的改進版。依靠以下性質：當最佳化凸光滑函數時，如果存在許多背離同一個點的下降方向，那它們之間任何的非負組合也會是下降方向。如我們在「懲罰函數的改進」小節中所示，我們使得 l_1 範數變得光滑，然後即可假設目標是光滑的。因此，我們提出將式 (27) 中這些 m 個步驟簡單相加的方法。因為每個步驟負責處理目標向量中的一個元素，我們可以將該加總寫成以下形式：

$$\begin{aligned} \mathbf{v}_0 &= \sum_{k=1}^m e_k \cdot \mathcal{S}_{\lambda/\|a_k\|} \left(\frac{a_k^T (\mathbf{x} - \mathbf{A}\mathbf{z}_0)}{\|a_k\|_2^2} + \mathbf{z}_0[k] \right) \\ &= \mathbf{S}_{W\lambda} (\mathbf{W}\mathbf{A}^T (\mathbf{x} - \mathbf{A}\mathbf{z}_0) + \mathbf{z}_0) \end{aligned} \quad (27)$$

其中 $\mathbf{W} = \text{diag}(\mathbf{A}^T \mathbf{A})^{-1}$ 。此項包含字典 \mathbf{A} 之各欄的範數。這些同時用於反向投影誤差 $\mathbf{A}^T (\mathbf{x} - \mathbf{A}\mathbf{z}_0)$ 的加權以及收縮運算元。該加權集合可以在演算法開始之前進行離線計算，而且有快速的方法可以逼近它們 [17]。請注意，得到的公式和式 (27) 不同，它不會要求從 \mathbf{A} 中提取欄，運算元需要的是直接與 \mathbf{A} 的乘積，以及其由向量和式 (23) 得到的伴隨矩陣。

雖然每一個 CD 方向都保證為下降，但如果缺乏適合標度，則它們的線性組合未必為下降。因此，如 [28] 所提到的，我們考慮此方向，並沿著它進行線性搜尋。這表示實際的反覆運算演算法如下：

$$\mathbf{z}_i = \mathbf{z}_0 + \mu (\mathbf{v}_0 - \mathbf{z}_0) \quad (28)$$

μ 是由某些類型的線性搜尋演算法所選擇。這就需要另一個與 \mathbf{A} 的乘積。這整個過程此後稱為 PCD 演算法。

相較於式 (23) 中的 SSF 演算法，PCD 的不同有兩個方面：1) \mathbf{A} 中元素的范數在後向投影誤差的加權中扮演重要角色，然而前面的演算法採用的是常量；2) 新演算法需要線性搜尋以得到下降方向。回想 SSF 中的常量 c 必須滿足 $c > \lambda_{\max}(\mathbf{A}^T \mathbf{A})$ 。另一方面，PCD 採用了與所使用加權矩陣 $\text{diag}(\mathbf{A}^T \mathbf{A})^{-1}$ 相同項的尺度值。如範例所示，對於依照歸一化矩陣 N 聯合的字典建立，SSF 需要 c 和 N ，這表示是式 (23) 中 $1/N$ 的一個權值。在 PCD 演算法中，由於 \mathbf{A} 中的欄是歸一化的，故加權矩陣是簡單恒等的。因此 PCD 花費 $O(N)$ 時間來增強 $\mathbf{A}^T (\mathbf{x} - \mathbf{A}\mathbf{z}_i)$ 項的加權，正因為如此，我們預計它的性能會更好。

上述收縮方法的其他說明也值得一提。在求解最小均方 (LS) 問題的領域中，Gauss-Seidel (GS) 演算法可以理解成連續的 CD 求解器。或者也可以選擇，對所有未背離當前解移動的座標，計算其座標的下降步數，然後將這些合併為一步。此結果與線性系統中著名的 Jacobi 方法一起，與此處提到的 PCD 平行發展。

加速技術

線性搜尋

上述演算法可以用很多方法進一步加速。首先，如同 PCD 描述般那樣執行線性搜尋的觀念，也與 SSF 演算法有關。這需要使用式 (23) 來計算一個臨時結果 \mathbf{v}_i ，然後再以

$\mathbf{z}_{i+1} = \mathbf{z}_i + \mu (\mathbf{v}_i - \mathbf{z}_i)$ 來定義解的形式，進行關於標量 μ 的 $f(\mathbf{z}_{i+1})$ 最佳化。

循序子空間最佳化

關於這些演算法第二種而且是更有效的加速選擇是部署 SESOP 方法 [44]。SESOP 的來歷源自於 CG 演算法 [45]。二次 CG (即 CG 應用於二次函數) 具有卓越的收斂性質。它的線性收斂速率 (參閱文獻 [36] 中的範例) 是 $\sqrt{r-1}/\sqrt{r+1}$ ，其中 r 是目標的 Hessian 條件個數。此速率比最快的下降速率 $(r-1)/(r+1)$ 要快得多。

也可以依賴二次 CG 的 $1/i^2$ 次線性最差案例收斂，其與 Hessian 調整無關 (參閱文獻 [47], [48] 和 [44] 中的範例)

$$f(\mathbf{z}_{i+1}) - f_{\text{optimal}} \leq \frac{L \|\mathbf{z}_0 - \mathbf{z}_{\text{optimal}}\|^2}{i^2} \quad (29)$$

其中 i 是反覆運算指標， L 是梯度 f 的 Lipschitz 常量。出現的收斂速率與二次 CG 眾所周知的擴展流型特性密切相關：在每一個反覆運算中，該方法在由所有先前的傳播步數和梯度張成的仿射子空間上對目標函數進行最小化。

在光滑凸函數的案例中 (不一定是二次的)，可以提出保留擴展流型特性的近似演算法。這樣的演算法可以在仿射子空間最佳化目標函數，該空間由所有先前傳播步數和最近梯度的方向張成。此方法繼承了 GC 的 $1/i^2$ 收斂性，然而，此方法的反覆運算成本將會隨著反覆運算的次數而增加。

為了緩解此問題，Nemirovski 在文獻 [47] 中建議將最佳化子空間限制在三個方向：當前梯度、所有先前步驟的加總，以及一個所有先前梯度的「巧妙」加權總和 (文獻 [44] 提出這種加權方法的一種版本)。所產生的 ORTH 方法繼承了最差情況下式 (29) 的最佳收斂 $1/i^2$ ，但當目標是二次時，它與 CG 並不一致，當函數在求解鄰域變成「近乎」二次時，它的收斂速度通常慢於 CG。

文獻 [44] 中的 SESOP 方法將 ORTH 子空間延伸到上次傳播步數的許多方向上。透過這種方式，當對於光滑凸函數保留 $1/i^2$ 收斂時，此方法變得與二次情況下的 CG 相當。此性質促進了演算法的效率。

在文獻 [17] 中，SESOP 與反覆運算收縮演算法相結合，由收縮步驟的方向取代當前的梯度方向。在二次情況下，所產生的 PCD-SESOP 方法相當於對角預處理的 CG 方法。在另一方面，PCD 和 SSF 的方向在初始步驟中提供比普通非線性 CG 更快的進度。這可以部分解釋這些方法在處理難題上的極致效能，這部分內容將會在後續的實驗章節中展示。

在每一個 SESOP 反覆運算上的低維子空間最佳化問題可以採用牛頓演算法來處理。在此過程中主要的計算負荷是需要與由 \mathbf{A} 生成的方向做乘積，但這些乘積可以儲存在先前的反覆運算中，使 SESOP 在幾乎沒有額外成本的情況下加速。在我們的經驗中，使用額外 ORTH 方向無法改善實際的收斂速度，因此，在我們的模擬中並沒採用它們。

NESTEROV 方法和 FISTA

在 1983 年，Nesterov 在文獻 [49] 中提到了一種方法，它對式 (29) 類型的函數具有 $1/i^2$ 收斂速率，可以作為 ORTH 的替代。Nesterov 的方法將梯度方向和先前反覆運算的結果以「巧妙」的方式給合，這不需要子空間最佳化或者線性搜尋。在另一方面，它通常慢於二次情況下的 CG 演算法，文獻 [44] 中提出 SESOP 優於 Nesterov 方法之處。

近來在文獻 [50] 中，Nesterov 將他的演算法推廣到包含兩個凸項的目標函數。其中一個是光滑的，它的梯度受限於 Lipschitz 常數，另一個則容易最佳化但為非光滑。這種情況類似於本文考慮過的 $l_2 - l_1$ 目標函數。採用 SSF 類型方法的步驟 (代替梯度)，結合先前反覆運算的結果，文獻 [50] 提出一種不需要線性搜尋的方法，同樣對式 (29) 具有 $1/i^2$ 的收斂速率，其中梯度的 Lipschitz 常數僅由光滑項決定。沿襲類似路線的還有一種叫做 FISTA 的替代方法，也獲得發展 [33]。簡而言之，FISTA 採用的反覆運算步驟與 SSF 十分類似，正如式 (22) 所示：

$$\mathbf{z}_{i+1} = \mathcal{S}_{\lambda/c} \left(\frac{1}{c} \mathbf{A}^T (\mathbf{x} - \mathbf{A} \mathbf{z}_i) + \mathbf{z}_i \right)$$

然而，替換掉應用此公式中的 \mathbf{z}_i ，一個如下的改進向量 $\hat{\mathbf{z}}_i$ 更適用

$$\hat{\mathbf{z}}_i = \mathbf{z}_i + \frac{t_i - 1}{t_i + 1} (\mathbf{z}_i - \mathbf{z}_{i-1})$$

其中 $\mathbf{z}_0 = 0$ ， $t_i = (1 + \sqrt{1 + 4t_{i-1}^2})/2$ 且 $t_1 = 1$ 。我們在後續的實驗章節中會對 FISTA 進行測試。

與共軛梯度結合

我們在此最後要提的加速選擇是一種泛化的 Polak-Ribiere CG 方法 [46]，其中可以結合 PCD 或 SSF 的方向。從尋找由 PCD (或 SSF，視情況而定) 獲得的方向開始，在第 i 個反覆運算時 $\mathbf{d}^0 = \mathbf{d}_{\text{PCD}}^0$ ，我們透過沿著 \mathbf{d}^i 方向進行準確的線性搜尋來計算 \mathbf{z}^{i+1} ，然後如式 $\mathbf{d}^{i+1} = \mathbf{d}_{\text{PCD}}^{i+1} + \beta_i \mathbf{d}^i$ 來準備下一個方向，其中

$$\beta_i = \frac{\nabla f(\mathbf{z}^{i+1})^T (\mathbf{d}_{\text{PCD}}^{i+1} - \mathbf{d}_{\text{PCD}}^i)}{\nabla f(\mathbf{z}^i)^T \mathbf{d}_{\text{PCD}}^i}$$

這裡 $\mathbf{d}_{\text{PCD}}^i$ 是一個在點 \mathbf{z}^i 上的 PCD 方向。我們稱這種方法為 PCD-CG。在二次目標函數最佳化的案例中，PCD-CG 相當於對角預處理的 CG 方法，因此它是快速漸近收斂。PCD-CG 和 SSF-CG 的全域收斂目前還沒有研究。

改進的懲罰函數

在我們轉向上述演算法的實驗演示之前，回到目標函數式 (3) 的定義並討論它的廣義化。在這個函數中，我們可以用廣義的懲罰函數代替 l_1 範數。

$$f_s(\mathbf{z}) = \frac{1}{2} \|\mathbf{x} - \mathbf{A} \mathbf{z}\|_2^2 + \lambda \sum_k \varphi_s(\mathbf{z}[k]) \quad (30)$$

其中應選擇 $\varphi_s(\tau)$ ，因為 1) 它提供更好的應用，2) 對於實現收縮運算的標量非線性特徵，它提供如式 (11) 的解析封閉運算式。

我們從 φ_s 的光滑和凸性選項開始，它易於加速收斂。雖然對這樣的函數來說，自然會用 $\varphi_s(\tau) = |\tau|^p$ ， $p > 1$ 來替代 $|\tau|$ ，但這不會導出標量非線性特徵的封閉運算式 [36]。式 [17] 提供具有吸引力的替代選擇

$$\varphi_s(\tau) = |\tau| - s \ln(1 + |\tau|/s), \quad s \in (0, \infty) \quad (31)$$

它能充分近似於 $|\tau|^p$ ，相當於一個取決於 s 的尺度因數。選擇 $\varphi(\tau)$ 的結果產生如下的非對稱標量非線性特徵

$$\mathcal{S}_{\lambda,s}(\tau_0) = \frac{|\tau_0| - \lambda - s + \sqrt{(|\tau_0| - \lambda - s)^2 + 4s|\tau_0|}}{2 \text{sign}(\tau_0)} \quad (32)$$

回到 $\varphi(\tau)$ 的凹性懲罰選項，一個如同 $|\tau|^p$ ， $0 < p < 1$ 的函數會導致大係數有較小失真，從而啓用「更精細」的最佳化，進一步將求解稀疏化，同時允許較大的非零係數輸入其中。此外，因為這個選項不允許標量非線性特徵的封閉運算式，我們可以利用對數函數

$$\varphi_s(\tau) = s \ln(1 + |\tau|/s), \quad s \in (0, \infty) \quad (33)$$

或者它的光滑替代項

$$\varphi_{cr}(\tau) = \frac{\ln(1 + |\tau|/s)}{\ln(1 + 1/s) - \frac{1}{p} \ln(1 + p/s)} - \frac{1}{p} \ln(1 + p|\tau|/s)$$

其中 p 和 s 控制原點的光滑性和凹度。這兩個函數都允許推導出封閉形式的標量非線性。

實驗研究

本章節中我們給出了四組實驗。首先是圖像去模糊，其次層析成像重建，第三是壓縮感測，第四是採取文獻 [51] 中的實驗，使用名為 geoGaussian 的合成矩陣 \mathbf{A} 。這些測試的共通處是需要將式 (3) 中的函數最小化，或者更確切地說，式 (31) 和 (32) 給出的光滑情形，這裡 $s = 0.001$ 。

在這些測試中要探究的演算法是 SSF [採用 $\mathbf{c} = \lambda_{\max}(\mathbf{A}^T \mathbf{A})$] 和 PCD、它們採用 CG 和 SESOP 的加速版 (採用七維，每個更新擁有多達七個牛頓步驟)、直接的 CG 和 SESOP 演算法、L-BFGS 演算法、FISTA [33] 和 L1-LS 內點法 [40]。對於 L-BFGS 演算法，我們利用 Mark Schmidt 寫的 MATLAB MINFUNC 包實現。

我們現在展示四個實驗及其結果。

圖像去模糊

追隨在「需要：稀疏和冗餘表示」小節中的說明，我們假設原圖像 \mathbf{x} (它們是理想圖像的實際像素) 經歷著名的去模糊運算 \mathbf{H} ，伴隨著方差為 σ^2 的加性雜訊。因此，我們將模糊和有噪圖像估量為 $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}$ ，目的是恢復 \mathbf{x} ，此時假定 \mathbf{x} 可以描述為 $\mathbf{A}\mathbf{z}$ ，其中帶有已知字典 \mathbf{A} 和稀疏向量 \mathbf{z} 。我們的目標是將以下式子最小化：

$$\hat{\mathbf{z}} = \arg \min_{\mathbf{z}} \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{A}\mathbf{z}\|_2^2 + \lambda \|\mathbf{z}\|_1$$

而且所恢復的圖像是 $\hat{\mathbf{x}} = \mathbf{A}\hat{\mathbf{z}}$ 。懲罰函數的組成如下：

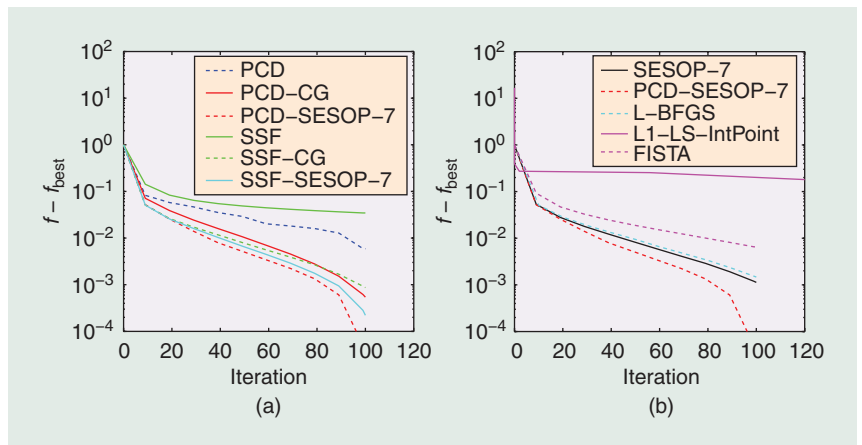
- 原圖像：我們的實驗採用大小為 128x128 像素的 Phantom 圖像，如圖3所示。
- 字典：我們使用 Daubechies 4 抽頭濾波器的冗餘 (非下採樣) 小波變換，分解層數為 4，冗餘度係數是 10 比 1
- 模糊核：按照文獻 [27] 中的測試，我們考慮一個 15x15 的核函數，在 $-7 \leq i, j \leq 7$ 時值為 $1/(i^2 + j^2 + 1)$ ，規範化到具有一個單位和。
- 雜訊性質：加性高斯白色雜訊是零均值，方差為 $\sigma = 0.01$ ，其中圖像的像素取值範圍為 $[0, 1]$ 。

圖 1-3 顯示在 $\lambda = 0.001$ 時得到的結果。此值是為達到滿意的結果而手動找到的。為了達到最佳性能，可以採用廣義的 Stein 無偏似然估計 (GSURE) 方法 (參閱參考文獻 [52] 和 [53]) 來調整此參數已獲得最佳效果。圖 1 顯示各種受測試演算法的目標值。如同所見，原始 PCD 和 SSF 演算法收斂較為緩慢，而它們的加速版本較快，事實上是這組方法中最快的。具有 SESOP 加速 (採用七維子空間) 的 PCD 和 SSF 其收斂速度比 CG、L-BFGS、L1-LS 及 FISTA 演算法還快。請注意，L1-LS 方法採用 inner-CG 步驟序列，因此圖中的標記之間存在比較寬的缺口。

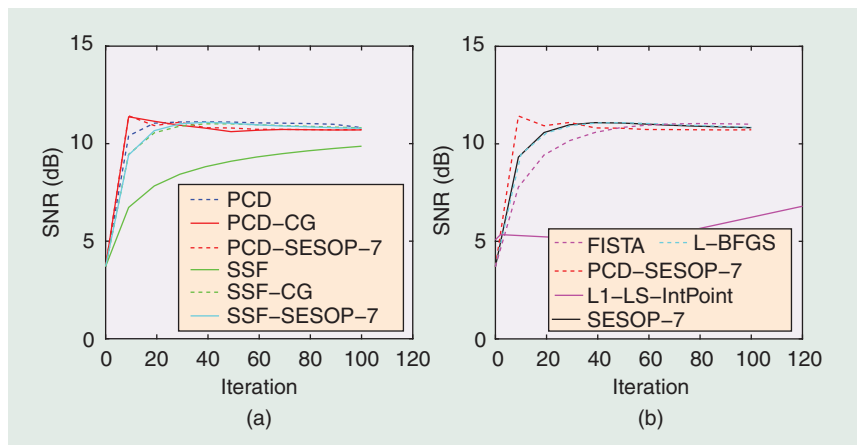
圖 2 達成我們實驗的真正目標——恢復原圖像。我們用信噪比 SNR 來衡量提議求解的品質：

$$\text{SNR} = 10 \cdot \log_{10} \left(\frac{\|\mathbf{x}\|_2^2}{\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2} \right) \quad (34)$$

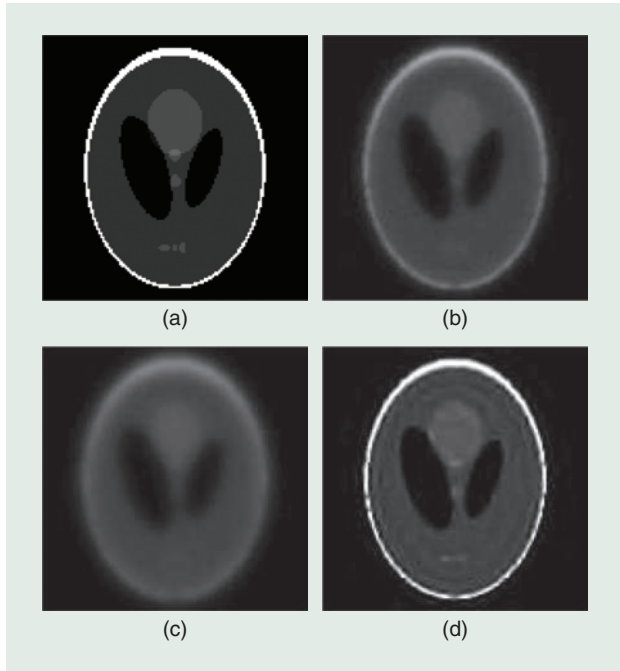
此值越高表示求解的品質越好。結果再次顯示加速 PCD (具有 SESOP-7 或 CG) 是最佳演算法，兩者都使用較少的反



【圖 1】去模糊：對不同測試演算法的反覆運算次數函數與目標函數值。



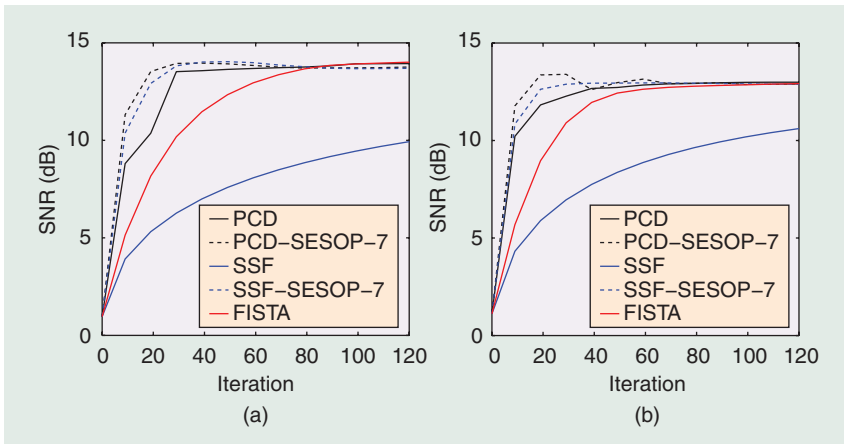
【圖 2】去模糊：對不同測試演算法的反覆運算次數函數與 SNR 值。



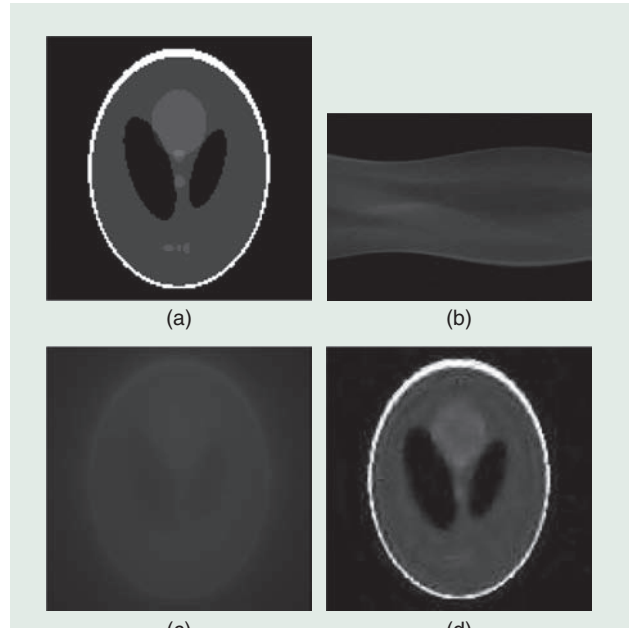
【圖 3】去模糊: (a) 原圖, (b) 觀測圖像, (c) 初始化圖像, (d) PCD-SESOP-7 演算法反覆運算 20 次後的恢復圖像。

覆運算獲得高品質的結果。由於當 SNR 達到尖峰時演算法會自動停止 [52], [53], 這意味著就輸出品質而言, 在本例中我們不需要重複超過十次的反覆運算。

圖 3 顯示作為測試對象的原圖像 Phantom。圖中還顯示觀測圖像 y 、初始化圖像 $A\hat{z}_0$, 以及 PCD-SESOP-7 反覆運算 20 次的結果。在所有的實驗中, 我們初始化 $\hat{z}_0 = \text{const} \cdot A^T H^T y$, 其中常數被選為用於適當歸一化之 HD 的最大奇異值。如同所見, 若結果受到一些人為影響, 其尖銳程度就會超過截至目前為止的其他測量。



【圖 5】X 光斷層攝影: 訊號維度的影響。(a) 曲線是表示 X 光斷層攝影技術應用於 32×32 像素的 Phantom 圖像。(b) 中曲線則對應著較高解析度的 256×256 像素圖像。



【圖 4】X 光斷層攝影: (a) 原圖像, (b) 觀測投影, (c) 初始化圖像, (d) PCD-SESOP-7 演算法 20 次反覆運算後的恢復圖像。

層析成像重建

對相同的圖像 Phantom 進行操作, 使用相同的字典, 並沿襲如上的相同公式, 現在我們轉向第二個實驗, 其中退化運算元 H 由 Radon 投影構成。該運算元產生 256 線投影, 每個包含 $\sqrt{2} \times 128 = 181$ 個標量, λ 如同先前選定 0.001。

圖 4 顯示所獲得的結果, 沿襲圖 3 中的相同結構。我們在圖中沒有顯示函數反覆運算時的目標值和 SNR, 因為它們與去模糊測試中的結果很類似 (參閱圖 1 和圖 2)。

圖 5 展示反覆運算收縮演算法的不同層面——它們的收斂趨勢對所處理訊號的維度來說很穩健。以相同的 X 光斷層攝影實驗為基礎, 我們進行兩個相似的測試, 一個實驗圖像的大小是 32×32 像素, 另一個圖像是大很多的 256×256 像素。圖 5 概述入選此組演算法的計算結果, 如圖所示, 收斂行為幾乎是一致的。這意味著恢復圖像需要相同的反覆運算次數, 而且與其維度無關。同時也說明, 這些重構過程的複雜度對於未知數仍為線性。

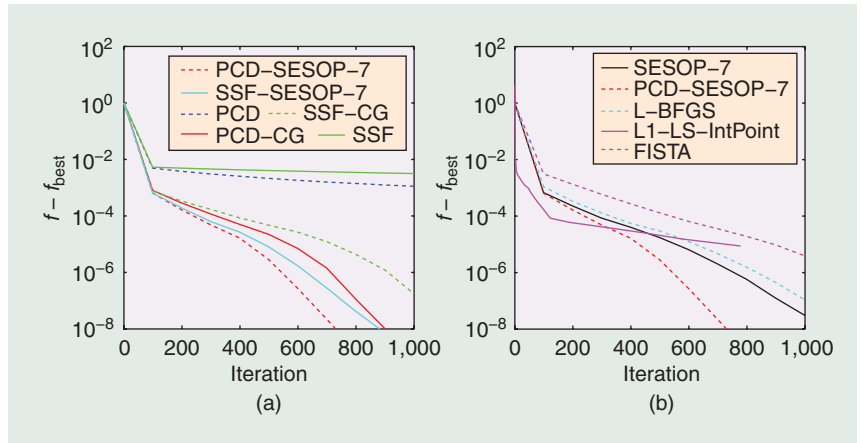
壓縮感測

第三個實驗是從圖像的隨機測量中恢復圖像，這也稱為壓縮感測。相對於先前的兩個實驗，我們對 64×64 像素的合成稀疏圖像 X 進行運算，只有 5% 的像素（隨機位置上）為非零值，出自獨立同分佈 (i.i.d.) 的高斯白色雜訊分佈 $\mathcal{N}(0, 10)$ 。這樣的圖像與天文或者雷達成像應用中所獲取的實際圖像相符。因為討論中的訊號是稀疏的，我們在公式中採用的字典是恒等運算元 $A = I$ 。

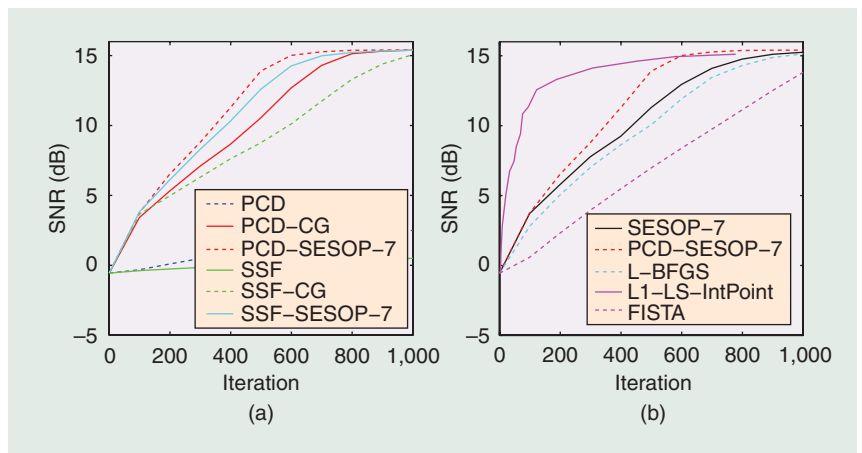
退化運算元 H (或 Q ，更符合「需要：稀疏和冗餘表示法」小節中的描述) 是由 410 個投影所構成，它們得自於隨機選擇對 X 套用二維快速傅葉爾轉換 (Fourier transform) 的 410 個樣本。因此， X 中的基數是 y 中測量值的兩倍，而我們的目標是恢復 x 。 y 中的值受到方差為 $\sigma = 0.01$ 的加性雜訊所污染。

圖 6 和圖 7 提出 $\lambda = 0.001$ 得出的結果，沿襲與先前實驗相同的結構。雖然大部分的演算法與之前的表現類似，L1-LS 卻表現出明顯差異，它在獲得高品質恢復結果方面快速有效。對比先前的兩個實驗，L1-LS 在此模擬中表現明顯不同的原因如下：在 z 中非零係數數目相對較少的情況下 (如本模擬)，預計 L1-LS 會十分有效，因為對數障礙函數的 Hessian 變成低秩矩陣，該秩對應於有效的約束條件數目 (訊號中的非零係數數目)。在本例中，用以解決牛頓系統的 inner-CG 方法變得十分有效。

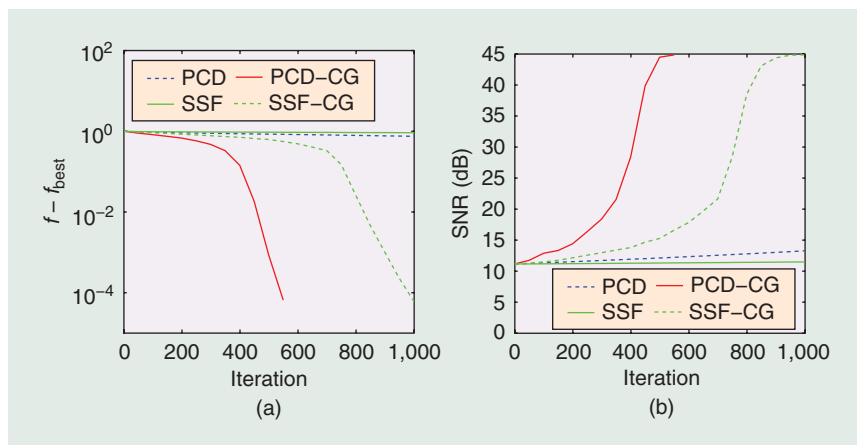
在目標函數中將 l_i 項替換為參數 $s = 1$ 的凹懲罰式(34)，可獲得大幅改善的結果。這樣會進一步促進結果的稀疏性，同更接近於理想圖像 X 。我們採用前一次用 PCD-SESOP-7 演算法進行 1000 次反覆運算的結果，並對同一組演算法進行初始化。圖 8 顯示精細化階段的目標函數和 SNR 值，透過 PCD-CG 演算法得到的結果具有明確且實質的改善。圖 9 顯示原始圖像、此



【圖 6】壓縮感測：對不同測試演算法的反覆運算次數函數與目標函數值。



【圖 7】壓縮感測：對不同測試演算法的反覆運算次數函數與 SNR 值。



【圖 8】壓縮感測：對於採用凹懲罰項的不同測試演算法，隨反覆運算次數函數變化的目標函數值和 SNR 值。

精細化階段的初始化與結果圖像，結果圖像看起來以近乎完美的方式重構指定圖像。

LORIS 資料的合成實驗

本文中第四個也是最後一個實驗使用文獻 [51] 中採取的實驗。此問題採用大小為 $1,848 \times 8,192$ 的顯式矩陣 \mathbf{H} ，稱之為 K_4 。此矩陣從高斯隨機矩陣建立而來，藉由塑造其奇異值來符合經典的高度病態地球科學問題（稱為 K_1 ）。至於壓縮感測案例，理想訊號 \mathbf{z}^* 以合成方法構建，其中向量長度為 8,192，輸入中的非零項占 5%（隨機位置），由獨立同分佈的高斯分佈 $\mathcal{N}(0, 10)$ 匯出。本例中的字典是單

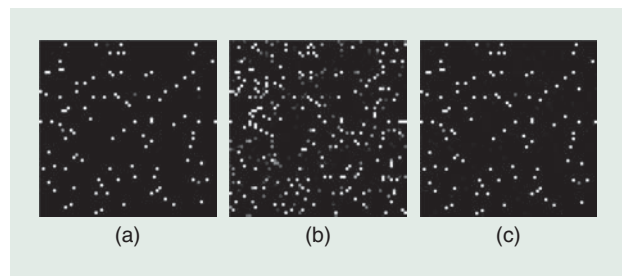
反覆運算收縮演算法構成全新的數值方法系列，以優於傳統最佳化方法的方式解決這些問題

位矩陣，我們的目標是從雜訊估計 $\mathbf{Hz}^* + \mathbf{v}$ 中恢復 \mathbf{z}^* ，其中 $\mathbf{v} \sim \mathcal{N}(0, \sigma \mathbf{I})$ 且 $\sigma = 1e - 4$ 。文獻 [51] 認為這是最難處理的問題。

除了比較不同演算法這個固有目標外，我們在此實驗中還要處理以下問題：1) 執行時間和相關的反覆運算次數，2) 演算法處理極小 λ 值時的表現，眾所皆知反覆運算搜尋演算法在這方面會退化。圖 10 顯示目標函數和 SNR，兩者分別輸入反覆運算計數和函數時間，且全部都使用 $\lambda = 1e - 3$ 。圖 11 則顯示套入 $\lambda = 1e - 6$ 的相同測試。

如我們所見，FISTA 的收斂速度比 SSF 快很多，這與文獻 [51] 中的觀察相呼應。在另一方面，PCD-SESOP 和 SSF-SESOP 遠比其他演算法更優異。在這個問題中沒有快速的矩陣乘積，因此反覆運算的代價主要受控於向量和 \mathbf{A} 之乘積，此外，單一 SESOP 的反覆運算速度與 FISTA 差不多。

我們也進行了文獻 [51] 中矩陣 K_2 和 K_3 的實驗（此處並未顯示），結果發現演算法之間的關係與上述結果相同。

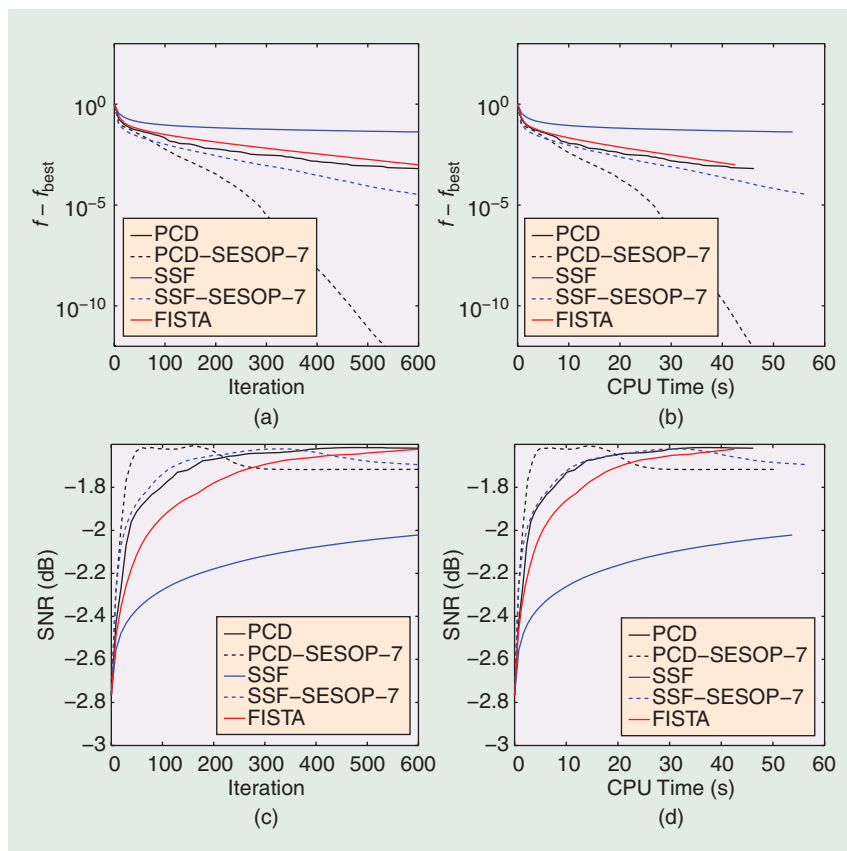


[圖 9] 壓縮感測：(a) 原圖像，(b) PCD-SESOP-7 反覆運算 20 次後所得到的圖像，(c) 採用凹目標函數和 PCD-CG 演算法 500 次反覆運算後的恢復圖像。

結論

就混合懲罰函數與稀疏求解的 $\ell_2 - \ell_1$ 最小化而言，反覆運算收縮演算法是很吸引人的系列方法，其效果足以超越傳統工具。在本文中，我們回顧這個領域的發展，介紹其中部分演算法，探索為這些方法進一步加速的方式，也研究這些演算法與其主要競爭者之間的表現比較。

本文進行一些實驗，以測試這些演算法的實際應用。這些實驗包括圖像趨模糊、層析成像重構、稀疏圖像的壓縮感測以及借自於文獻 [51] 的一個測試，此測試模仿地球物理方面的高度病態逆問題。這些實驗均顯示反覆運算收縮方法的確十分有效，尤其是合併 SESOP 加速之後。在我們研究的所有演算法中，PCD-SESOP 在大部的測試中表現最優。FISTA 是最近才出現的加速演算法選擇，在運算元 \mathbf{A} 為隱式時，速度可能比 SESOP 方法還快，而且它還具有快速部署演算法。



[圖 10] K4 實驗：採用 $\lambda = 1e - 3$

針對這類型演算法所做的研究明顯不足，尚待努力的部分包括提出新的演算法、求解各種還未解決的問題，以及找出可從受惠這些工具的有趣應用。我們希望本文能夠激勵讀者進一步研究這些課題，以拓展此凸最佳化特殊分支的知識和功能。為了支援這類活動，我們發佈了可以重現本文所有圖文的完整的 MATLAB 工具包。此工具包可免費使用，請從網址 <http://ie.technion.ac.il/~mcib/> 下載名為 iterative-shrinkage package 的檔案。

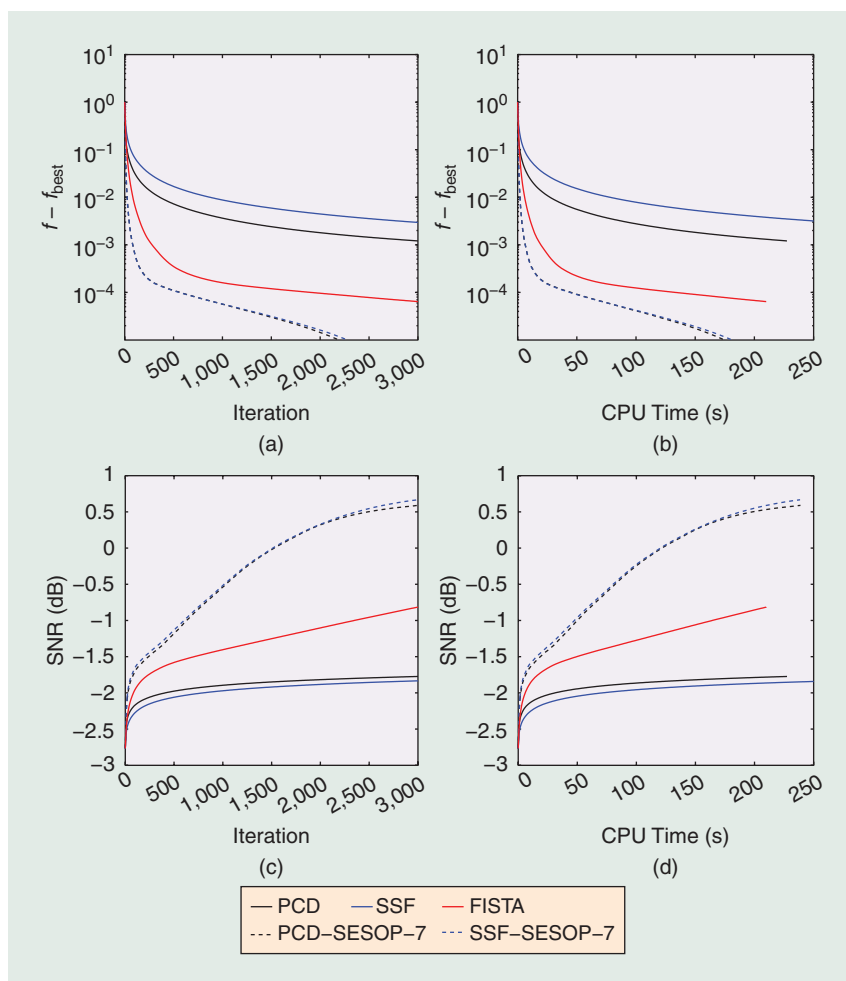
致謝

This research was partly supported by the Israel Science Foundation grant number 1031/08, and by the European Community's FP7-FET program, SMALL project, under grant agreement number 225913. We would like to thank Mario Figueiredo and Ignace Loris for sharing their software packages with us. We are also thankful to Michael Lustig for fruitful discussions on the various algorithms explored.

作者簡介

Michael Zibulevsky (mzib@cs.technion.ac.il) received his M.Sc. degree in electrical engineering from MIIT, Moscow, and the D.Sc. degree in operations research (non-linear optimization) from Technion, Israel, in 1996. He spent two years with the University of New Mexico working on blind source separation via sparse signal representations, forming the foundation to this field. He is with the Department of Computer Science at Technion. His area of interests include non-linear optimization, signal/image processing, computer tomography, ultrasound imaging, and EEG/MEG. He is also active in the field of sparse approximation and its application to image processing.

Michael Elad (elad@cs.technion.ac.il) received his B.Sc., M.Sc., and D.Sc. degrees in 1986, 1988, and 1997, respectively, from the Department of Electrical Engineering at Technion, Israel. Since 2003, he has been a faculty member in the Computer Science Department at Technion. He works in the field of signal and image processing, specializing in particular on sparse representations and over-complete transforms. He received Technion's Best Lecturer Award six times, the 2007 Solomon Simon Mani Award for excellence in teaching, and the 2008 Henri Taub Prize for academic excellence. He is an associate editor for *IEEE Transactions on Image Processing* and *SIAM Journal on Imaging Sciences*. He is a Senior Member of IEEE.



【圖 11】K4 實驗：採用 $\lambda = 1e - 6$

參考文獻

- [1] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, Feb. 2009.
- [2] S. Mallat, *A Wavelet Tour of Signal Processing—The Sparse Way*, 3rd ed. Burlington, MA: Academic, 2009.
- [3] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.
- [4] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. R. Statist. Soc. Ser. B*, vol. 58, no. 1, pp. 267–288, 1996.
- [5] D. L. Donoho and M. Elad, "Optimally sparse representation in general (non-orthogonal) dictionaries via l_1 minimization," *Proc. Nat. Acad. Sci.*, vol. 100, no. 5, pp. 2197–2202, 2003.
- [6] D. L. Donoho and M. Elad, "On the stability of the basis pursuit in the presence of noise," *Signal Process.*, vol. 86, no. 3, pp. 511–532, Mar. 2006.
- [7] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
- [8] J. A. Tropp, "Just relax: Convex programming methods for subset selection and sparse approximation," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 1030–1051, Mar. 2006.
- [9] J.-L. Starck, E. J. Candès, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Image Processing*, vol. 11, no. 6, pp. 670–684, 2002.
- [10] J. Bioucas-Dias, "Bayesian wavelet-based image deconvolution: A GEM algorithm exploiting a class of heavy-tailed priors," *IEEE Trans. Image Processing*, vol. 15, no. 4, pp. 937–951, Apr. 2006.
- [11] M. J. Fadili and J. L. Starck, "Sparse representation-based image deconvolution by iterative thresholding," in *Proc. Astronomical Data Analysis ADA' 06*, Marseille, France, Sept. 2006.

- [12] M. Elad, J.-L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *J. Appl. Comput. Harmon. Anal.*, vol. 19, no. 3, pp. 340–358, 2005.
- [13] J. Bobin, J.-L. Starck, J. Fadili, Y. Moudden, and D. L. Donoho, "Morphological component analysis: An adaptive thresholding strategy," *IEEE Trans. Image Processing*, vol. 16, no. 11, pp. 2675–2681, 2007.
- [14] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [15] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [16] M. Lustig, D. L. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, Dec. 2007.
- [17] M. Elad, B. Matalon, and M. Zibulevsky, "Coordinate and subspace optimization methods for linear least squares with non-quadratic regularization," *Appl. Comput. Harmon. Anal.*, vol. 23, pp. 346–367, Nov. 2007.
- [18] B. Efron, T. Hastie, I. M. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Statist.*, vol. 32, no. 2, pp. 407–499, 2004.
- [19] Y. Tsaig and D. L. Donoho, "Fast solution of ℓ_1 -norm minimization problems when the solution may be sparse," *IEEE Trans. Inform. Theory*, vol. 54, no. 11, pp. 4789–4812, Nov. 2008.
- [20] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [21] P. A. Jansson, *Deconvolution: With Applications in Spectroscopy*. New York: Academic, 1984.
- [22] J. C. Hoch, A. S. Stern, D. L. Donoho, and I. M. Johnstone, "Maximum-entropy reconstruction of complex (phase-sensitive) spectra," *J. Magn. Reson.*, vol. 86, no. 2, pp. 236–246, Feb. 1990.
- [23] J.-L. Starck, F. Murtagh, and A. Bijaoui, "Multiresolution support applied to image filtering and restoration," *Graph. Models Image Processing*, vol. 57, no. 5, pp. 420–431, Sept. 1995.
- [24] N. G. Kingsbury and T. H. Reeves, "Overcomplete image coding using iterative projection-based noise shaping," in *Proc. Int. Conf. Image Processing (ICIP)*, 2002, vol. 3, pp. 597–600.
- [25] M. A. Figueiredo and R. D. Nowak, "An EM algorithm for wavelet-based image restoration," *IEEE Trans. Image Processing*, vol. 12, no. 8, pp. 906–916, 2003.
- [26] I. Daubechies, M. DeFrise, and C. De-Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, 2004.
- [27] M. A. Figueiredo and R. D. Nowak, "A bound optimization approach to wavelet-based image deconvolution," in *Proc. IEEE Int. Conf. Image Processing—ICIP 2005*, Genoa, Italy, Sept. 2005, vol. 2, pp. 782–785.
- [28] M. Elad, "Why simple shrinkage is still relevant for redundant representations?," *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5559–5569, Dec. 2006.
- [29] T. Adeyemi and M. E. Davies, "Sparse representations of images using overcomplete complex wavelets," in *Proc. IEEE SP 13th Workshop Statistical Signal Processing*, Bordeaux, France, July 17–20, 2006, pp. 805–809.
- [30] M. A. Figueiredo, J. M. Bioucas-Dias, and R. D. Nowak, "Majorization-minimization algorithms for wavelet-based image restoration," *IEEE Trans. Image Processing*, vol. 16, no. 12, pp. 2980–2991, 2007.
- [31] T. Blumensath and M. E. Davies, "Iterative thresholding for sparse approximations," *J. Fourier Anal. Appl.*, vol. 14, no. 5–6, pp. 629–654, Dec. 2008.
- [32] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for ℓ_1 -minimization with applications to compressed sensing," *SIAM J. Imaging Sci.*, vol. 1, no. 1, pp. 143–168, 2008.
- [33] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, 2009.
- [34] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [35] E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," in *Proc. Int. Conf. Image Processing*, Lausanne, Switzerland, Sept. 1996, pp. 379–383.
- [36] P. Moulin and J. Liu, "Analysis of multiresolution image denoising schemes using generalized Gaussian and complexity priors," *IEEE Trans. Inform. Theory*, vol. 45, no. 3, pp. 909–919, 1999.
- [37] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting," *SIAM J. Multiscale Model. Simul.*, vol. 4, no. 4, pp. 1168–1200, Nov. 2005.
- [38] S. Sardy, A. G. Bruce, and P. Tseng, "Block coordinate relaxation methods for nonparametric signal denoising with wavelet dictionaries," *J. Comput. Graph. Statist.*, vol. 9, no. 2, pp. 361–379, 2000.
- [39] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmon. Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.
- [40] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "A Method for large-scale ℓ_1 -regularized least squares problems with applications in signal processing and statistics," *IEEE J. Select. Topics Signal Processing*, vol. 1, no. 4, pp. 606–617, Dec. 2007.
- [41] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM J. Control Optim.*, vol. 14, no. 5, pp. 877–898, 1976.
- [42] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood estimation from incomplete data via the EM algorithm," *J. R. Statist. Soc. Ser. B*, vol. 39, no. 1, pp. 1–38, 1977.
- [43] K. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions (with discussion)," *J. Comput. Graph. Statist.*, vol. 9, no. 1, pp. 1–59, 2000.
- [44] G. Narkiss and M. Zibulevsky, "Sequential subspace optimization method for large-scale unconstrained optimization," *Technion, Israel Inst. Technol.*, Haifa, Tech. Rep. CCIT 559, 2005.
- [45] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *J. Res. Natl. Bur. Stand.*, vol. 49, no. 6, pp. 409–436, 1952.
- [46] J. Nocedal and S. Wright, *Numerical Optimization*. New York: Springer, 1999.
- [47] A. Nemirovskii, "Orth-method for smooth convex optimization," (in Russian), *Izvestia AN SSSR, Ser. Tekhnicheskaya Kibernetika*, vol. 2, 1982 (Transl.: Eng. Cybern. Soviet J. Comput. Syst. Sci.).
- [48] A. S. Nemirovskii and D. B. Yudin, *Problem Complexity and Method Efficiency*. New York: Wiley, 1983.
- [49] Y. E. Nesterov, "A method for solving the convex programming problem with convergence rate $O(1/k^2)$," (in Russian), *Dokl. Akad. Nauk SSSR*, vol. 269, pp. 543–547, 1983.
- [50] Y. E. Nesterov. (2007). *Gradient methods for minimizing composite objective function*. CORE Report [Online]. Available: http://www.ecore.be/DPs/dp_1191313936.pdf
- [51] I. Loris, "On the performance of algorithms for the minimization of ℓ_1 -penalized functionals," *Inverse Probl.*, vol. 25, no. 3, pp. 1–16, 2009.
- [52] C. Vonesch, S. Ramani, and M. Unser, "Recursive risk estimation for non-linear image deconvolution with a wavelet-domain sparsity constraint," in *Proc. 15th IEEE Int. Conf. Image Processing (ICIP)*, 2008, pp. 665–668.
- [53] R. Giryes, M. Elad, and Y. C. Eldar, "The projected GSURE for automatic parameter tuning in iterative shrinkage methods," submitted for publication.



主編鄧立
IEEE 信號處理學會發行經理
Linda Cherry
deng@microsoft.com
<http://signalprocessingsociety.org/publications/periodicals/spm>

訊號處理雜誌之全球影響

Thomson-Reuters ISI 每年都發佈期刊引證報告 (JCR)，透過確定期刊的文章是否經常被後來之研究所引用，審查科技期刊之影響。

連續兩年，**IEEE 訊號處理雜誌 (Signal Processing Magazine, SPM)** 高居全球所有電氣和電子工程類 247 種出版物及所有 127 種 IEEE 出版物榜首。

我們致力於出版頗具影響力之高品質文章。在 2010 的 ISI 報告中，SPM 的影響指數為 5.86，相較 2009 年的影響指數 4.91 有明顯提升。訊號處理之社會影響，及其非凡的各種應用構成了我們雜誌最高影響指數的潛在基礎。SPM 與 IEEE 的影響深遠。IEEE

的科學資訊推動了許多的技術創新。IEEE 被引用的科技和技術文章往往是其他任何學術或商業出版商的三倍。

我們雜誌的特點是其以適當的數學深度和嚴謹性廣泛報導不同的技術領域。其目標是在宣導發展統一方法的同時，有助於組織涉獵廣泛的訊號處理主題。許多卓有成效的訊號處理技術都很成功，不限於一個應用領域，經常也出現在其他領域，儘管往往需要極具見地的應用針對特定領域的直覺。我們雜誌的主要目的是提供相互影響的不同應用領域間，所用之相同方法和精確處置的指導性文章。我們所有的特刊都是由頂尖編輯團隊運作。

我們的編輯一直積極地拓展 SPM 之服務範圍。透過電子郵件遞送雜誌的數位版給世界各地讀者，讓我們的會員能夠更便捷地瞭解尖端研究和創新產業之發展趨勢。我們的翻譯將為更多的學生讀者和從業著讀者提供創新之科技資訊。

作為訊號處理社群，讓我們擁抱我們所在領域的生命力，慶祝我們的工作所具有的全球影響力，并分享引用影響獨佔鰲頭之榮光。此榮光體現了我們工作的品質和所在領域的社會影響力。

SPM

IEEE SIGNAL PROCESSING MAGAZINE

Li Deng, Editor-in-Chief — Microsoft Research

AREA EDITORS

Feature Articles — Antonio Ortega, University of Southern California
Columns and Forums — Ghasan AlRegib, Georgia Institute of Technology
Special Issues — Dan Schonfeld, University of Illinois at Chicago
e-Newsletter — Z. Jane Wang, University of British Columbia

EDITORIAL BOARD

Les Atlas — University of Washington
Jeff Bilmes — University of Washington
Holger Boche — Technische Universität München
Yen-Kuang Cheng — Intel Corporation
Liang-Gee Chen — National Taiwan University
Ed Delp — Purdue University
Adriana Dumitras — Apple Inc.
Brendan Frey — University of Toronto
Alex Gershman — Darmstadt University of Technology, Germany
Mazin Gilbert — AT&T Research
Bernd Girod — Stanford University
Jenq-Neng Hwang — University of Washington
Michael Jordan — University of California, Berkeley
Vikram Krishnamurthy — University of British Columbia, Canada
Chin-Hui Lee — Georgia Institute of Technology
Jian Li — University of Florida-Gainesville

Mark Liao — National Chiao-Tung University, Taiwan
Hongwei Liu — Xidian University, China
K.J. Ray Liu — University of Maryland
Tom Luo — University of Minnesota
Nelson Morgan — ICSI and University of California, Berkeley
Fernando Pereira — ISTIT, Portugal
Roberto Pieraccini — Speech Cycle Inc.
H. Vincent Poor — Princeton University
Nicholas Sidiropoulos — Tech University of Crete, Greece
Yoram Singer — Google Research
Henry Tirri — Nokia Research Center
Anthony Vetro — MERL
Patrick J. Wolfe — Harvard University

ASSOCIATE EDITORS— COLUMNS AND FORUM

Andrea Cavallaro — Queen Mary, University of London
Rodrigo Capobianco Guido — University of Sao Paulo, Brazil
Deepa Kundur — Texas A&M
Andres Kwasinski — Rochester Institute of Technology
Rick Lyons — Besser Associates
Aleksandra Mojsilovic — IBM T.J. Watson Research Center
Douglas O'Shaughnessy — INRS, Canada
Greg Slabaugh — Medicsight PLC, U.K.
Clay Turner — Pace-O-Matic, Inc.
Alessandro Vinciarelli — IDIAP-EPFL
Stephen T.C. Wong — Methodist Hospital-Cornell
Dong Yu — Microsoft Research

ASSOCIATE EDITORS—E-NEWSLETTER

Marcelo Bruno — ITA, Brazil
Gwenaél Doerr — Technicolor, France
Shantanu Rane — MERL
Yan Lindsay Sun — University of Rhode Island

IEEE PERIODICALS MAGAZINES DEPARTMENT

Geraldine Krolin-Taylor — Senior Managing Editor
Jessica Barragué — Managing Editor
Susan Schneiderman — Business Development Mgr.
+1 732 562 3946 Fax: +1 732 981 1855
Felicia Spagnoli — Advertising Production Mgr.
Janet Dudar — Senior Art Director
Gail A. Schnitzer — Assistant Art Director
Theresa L. Smith — Production Coordinator
Dawn M. Melley — Editorial Director
Peter M. Tuohy — Production Director
Fran Zappulla — Staff Director, Publishing Operations

IEEE prohibits discrimination, harassment, and bullying.
For more information, visit <http://www.ieee.org/web/aboutus/whatis/policies/p9-26.html>.

IEEE SIGNAL PROCESSING SOCIETY

Mos Kaveh — President
K.J. Ray Liu — President-Elect
Michael D. Zoltowski — Vice President, Awards and Membership
V. John Mathews — Vice President, Conferences
Min Wu — Vice President, Finance
Ali H. Sayed — Vice President, Publications
Ahmed Tewfik — Vice President, Technical Directions
Mercy Kowalczyk — Executive Director and Associate Editor
Linda C. Cherry — Manager, Publications

GLOSSARY OF TERMS

Academic terms used in Mainland China and Hong Kong, Taiwan

English	Mainland China	Hong Kong, Taiwan
ADC	模拟/数字转换器	類比/數位轉換器
Bayesian	贝叶斯	貝氏
chip	芯片	晶片
clustering	聚类	群集
compressed sensing	压缩感知	壓縮感測
computer code	计算机代码	電腦程式碼
computer vision	计算机视觉	電腦視覺
data	数据	資料
digital	数字	數位
factor	因子	因數
global minimum	全局最小	全域最小
hierarchical	分层的	階層
image	图像	影像
intellectual property	知识产权	智慧財產權
iterative	迭代	反覆運算
labeling	标记	標貼
least square	最小二乘	最小平方
linear approximation	线性近似	線性模型擬合
module	模块	模組
Monte Carlo	蒙特卡罗	蒙地卡羅
noise	噪声	雜訊
operator	算子	運算子
optimization	最优化	最佳化
Physical layer	物理层	實體層
priori	先验的	事先
probability	概率	機率
program	程序	程式
read-only memory	只读存储器	唯讀記憶體
real time	实时	即時
robustness	鲁棒性	穩健性
sampling	采样	抽樣
sensor	传感器	感測器
signal	信号	訊號
software	软件	軟體
system-on-Chip	系统级芯片	系統單晶片
transistor	晶体管	電晶體
variable	变量	變數
variance	方差	變異數
video	视频	視訊

Can new telemedicine standards keep us healthier longer?

Find the latest research in IEEE *Xplore*

Wherever you find people developing the most advanced technology, chances are you'll find them using the IEEE *Xplore* digital library. That's because IEEE *Xplore* is filled with the latest research on everything from wireless technology and solid-state circuits—to telemedicine standards that can help us get more out of life.

When it comes to technology, the research that matters is in IEEE *Xplore*.



See for yourself. Read "Evaluation of Security and Privacy Issues in Integrated Mobile Telemedical System," only in IEEE *Xplore*.

**Try IEEE *Xplore* free—
visit www.ieee.org/livehealthier**

IEEE *Xplore*[®] Digital Library
Information driving innovation





了解详情:

www.mathworks.cn/accelerate

产品资料

成功案例

申请试用软件

使用PARALLEL COMPUTING TOOLBOX™ 并行化 MATLAB® 应用程序

如果您知道如何写FOR循环,您就知道如何写一个MATLAB并行程序。Parallel Computing Toolbox的任务并行和数据并行程序功能,让您能够快速在多核计算机或者集群上运行MATLAB程序。

MATLAB®
& SIMULINK®

