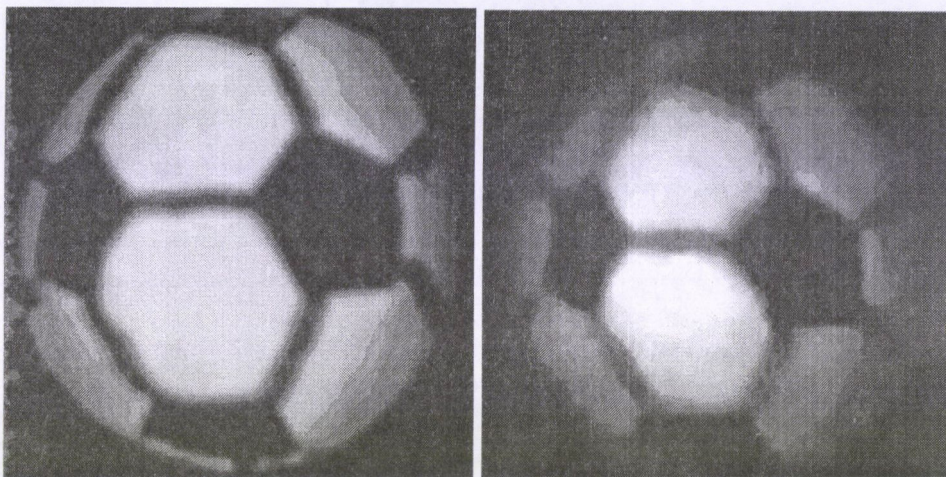


TÖMÖRÍTETT ÉRZÉKELÉSNEEL MINDEN EGYES KÉPPONT SZÁMÍT¹

Fordította: Réti Zoltán

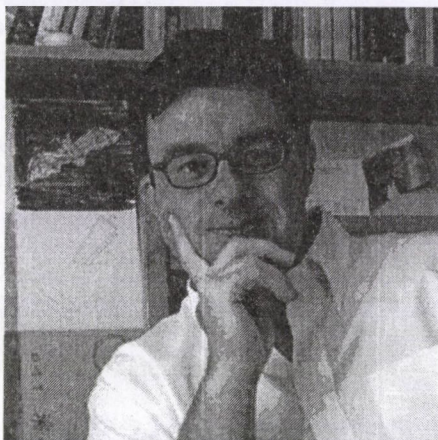


Egy is elegendő. Richard Baraniuk és Kevin Kelly a Rice University-ről építette azt az egyetlen pixeles fényképezőgépet, amivel ez a fénykép készült. (a) Egy focilabda képe, amit 64x64-es felbontásban hagyományos digitális fényképezőgép vett le. (b) Ugyanaz a focilabda az egyetlen pixeles fényképezőgéppel. A képet 1600 különböző, véletlenszerűen kiválasztott mérésből matematikailag származtatták, felhasználva a tömörített érzékelésnek nevezett módszert. (R. G. Baraniuk, Compressive Sensing [Lecture Notes], Signal Processing Magazine, July 2007. c2007 IEEE. szívességéből)

Az a közös dolog a szemétnben és a számítógépes fájlokban, hogy mindkettő összenyomottan szép. A digitális fényképezőgépek vásárlásakor azonban észrevehetjük, hogy gyártóik ezt nem értették meg.

Néhány éve a műszaki boltok tele voltak 1, vagy 2 megapixeles gépekkel. Aztán jöttek a 3 megapixeles áramköri lapkák, majd a 10, sőt a 60 megapixelesek.

¹Eredeti cím: Compressed Sensing Makes Every Pixel Count, What's Happening in the Mathematical Science, 7. kötet. Az Alkalmazott Matematikai Lapok szerkesztősége köszönetet mond az Amerikai Matematikai Társaságnak (American Mathematical Society), amiért engedélyezte a dolgozat magyar fordításának egyszeri publikálását. Az AMS minden további jogot fenntart magának.

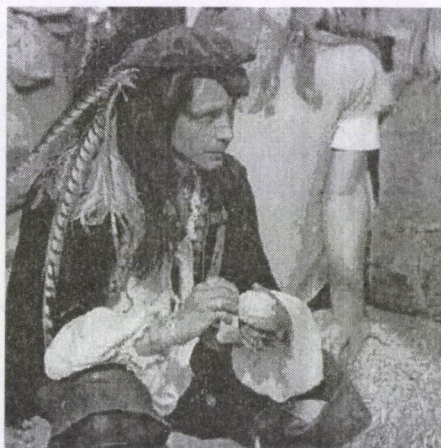
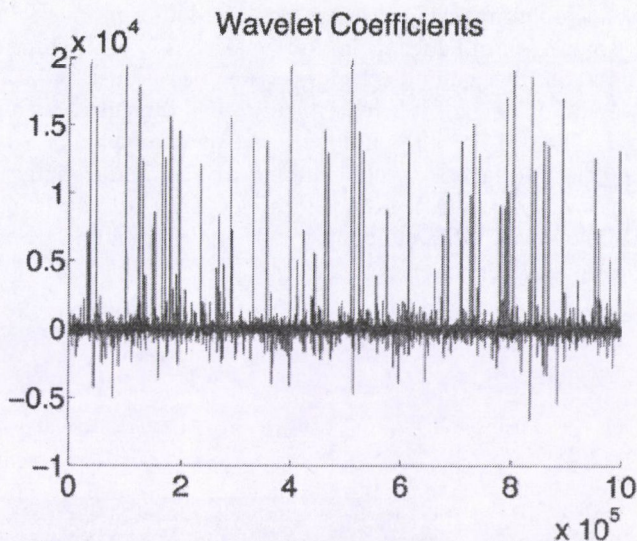


Emmanuel Candes. (saját szívésségéből)

Sajnos, ezek a nagyon sok pixeles fényképezőgépek óriási méretű számítógépes adatfájlokat hoznak létre. Ezért ha az ember e-mail-en akar elküldeni egy fényképet, vagy fel akarja rakni a világhálóra, az az első dolog, hogy kezelhetőbb méretre tömöríti. Általában szabad szemmel lehetetlen észrevenni a különbséget ezen összetömörített és az eredeti fénykép között (lásd az 1. ábrát). Így kettős dinamika fejlődött ki: a fényképezőgépek mérnökei egyre több adatot zsúfolnak a lapkákra, miközben a számítógépes programok mérnökei egyre okosabb módokat ötlenek ki, hogy megszabaduljanak az adatoktól.

2004-ben matematikusok megtalálták a módját, hogy hogyan állítsák meg ezt a „fegyverkezési versenyt”. Miért vegyünk méretet 10 milliószor, amikor esetleg 10 ezer is elég ahhoz, hogy képünket kielégítően jellemezze? Nem lenne-e jobb, ha már kezdetben is csak a leglényegesebb 10 ezer adatot vennénk le? Hálával tartozunk Emmanuel Candesnek (Caltech Egyetem), Terence Taonak (University of California Los Angeles-i részlege), Justin Rombergnek (Georgia Tech Egyetem) és David Donohonak (Stanford University) azért a hathatós matematikai technikáért, amivel az adatmennyiséget ezredére csökkenthetjük, *mielőtt* egyáltalán befogadnánk. Eljárásukat *tömörített érzékelésnek* hívják. Ez a kifejezés mérnököknél mostanában nagyon felkapott lett, habár matematikai gyökerei évtizedesek.

Az elgondolás bizonyítékeként Richard Baraniuk és Kevin Kelly (Rice University) még egy *egyetlen pixeles* fényképezőgépet is kifejlesztettek. Azért persze ne gondoljuk, hogy ez a gép felbukkan a Wal-Mart polcain a 10 megapixelesek mellett, mivel utóbbiaknak beépített gazdaságossági előnyük van. „Azért tudjuk olyan olcsón megépíteni őket, mert szerencsés véletlen egybeesés van a szemünk által és a szilikon által érzékelt fény hullámhosszai között.” - mondja Baraniuk. „Ez tette lehetővé, hogy a fényképezőgép-készítők a Moore-szabály szekerén utazzanak (azaz kihasználják Moore-szabályát)”, más szóval azt, hogy nagyjából kétszerként megduplázzák a pixeleket, azaz képpontok számát.

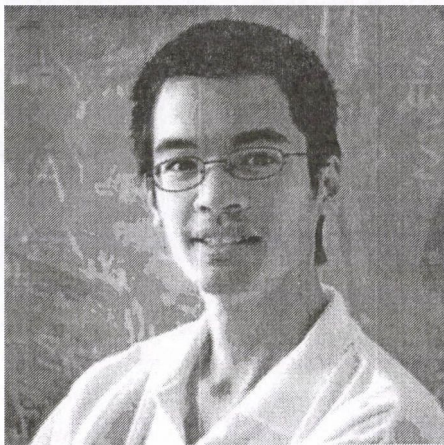


1. ábra. Balra: az eredeti teszt kép. Fent: Az egyik standard tömörítő eljárás, hogy a képet waveletek összegeként fejezzük ki. Fent a waveletek együtthatóit ábrázoltuk, a nagy együtthatók a képhez jelentősen hozzájáruló waveleteket jelzik (például amelyek egy élet vagy textúrát azonosítanak). Jobbra: Amikor a kis együtthatókkal rendelkező waveleteket kihagyjuk, és csak a többi waveletből állítjuk a képet helyre, az eredetitől majdnem megkülönböztethetelen lesz. (Emmanuel Candes szívességéből)

Emiatt a tömörített érzékelés valódi piaca a láthatatlan hullámhosszoknál van. Ezen hullámhosszoknál nem olcsó dolog érzékelőket építeni, viszont az ilyen érzékelőknek sok alkalmazási lehetősége van. Például a mobiltelefonok a rádiófrekvenciák széles spektrumából fognak kódolt jeleket. A terahertzes sugárzás² vevőit pedig csempészáru vagy ruha alatt elrejtett fegyverek kiszűrésére lehet használni.

Még a közönséges infravörös fényről is költséges fényképet készíteni. „Ha kilépünk abból a hatókörből, ahol a szilikon érzékeny, akkor a 100 dolláros fényképezőgép 100 ezer dollárossá válik” - mondja Baraniuk. Néhány alkalmazásnál, például úrhajók esetében, lehet, hogy nincs elég hely túl sok érzékelő számára. Az ilyen alkalmazásoknál komolyan el kell gondolkodnunk azon, hogy hogyan tegyünk minden egyes képpontot (pixel) értékké.

1. A régi, hagyományos bölcsesség



Terence Tao. (Reed Hutchinson/UCLA szívességéből)

A tömörített érzékelés története Claude Shannonnal kezdődik, aki az információelmélet úttörője volt. Shannon 1949-ben bizonyította be, hogy egy időben változó, de N hertznél nem magasabb frekvenciájú jelet tökéletesen helyre lehet állítani, ha a jelből $\frac{1}{2}N$ másodperces szabályos időközönként mintát veszünk. De ez pont annak a tételnek a fordítottja, ami jelfeldolgozók nemzedékeinek vált az evangéliumává: az N hertznél magasabb frekvenciájú jelet *lehetetlen* egyértelműen helyreállítani; egy jel bármikor „álméven” jelenhet meg (aliasing), vagyis mintája azonos lehet egy másik jelből vett mintával.

²Ez az elektromágneses spektrum azon része, amit nézőponttól függően vagy ultra ultramagas rádiófrekvenciának, vagy infra infravörös fénynek írhatunk le.

A digitális képek világában egy kép a „jel”, és a képből vett „minta” jellegzetesen egy képpont (pixel), ami más szóval egy kiválasztott pont mérésnyi fényerőssége (talán színadatokkal is megtűzdelve). Shannon tétele, amit Shannon-Nyquist mintavételi tételnek is neveznek, azt mondja ki, hogy egy kép felbontása egyenesen arányos a mérések számával. Ha meg akarjuk duplázni a felbontást, akkor jól tesszük, ha a pixelek számát is megduplázzuk. A digitális fényképezőgépek eladói pontosan ilyennek látják a világot.

Candes, Tao, Romberg és Donoho ezt a világot állította a feje tetejére. A tömörített érzékelés világában az elérhető felbontást főleg a kép *információtartalma* vezérli. Alacsony információtartalmú képet tökéletesen helyre lehet állítani kis számú mérésből. Ha már megfelelő számú mérést végeztünk, nem segít, ha még többet mérünk. Ha az ilyen képek ritkák vagy szokatlanok lennének, a dolog nem lenne nagyon izgalmas. De valójában *jóformán mindegyik*, a valós életből vett képnek alacsony az információtartalma (mint ahogy az 1. ábrán látható).

Ez a kijelentés az ösztöneinkkel ellentétesnek tűnik, hiszen az „információ” matematikai jelentése a köznyelvi jelentésnek a majdnem teljes ellentéte. Egy TV-képernyőn levő véletlenszerű statikus zaj látványa példa magas információtartalmú képre. A legtöbb laikus valószínűleg ezt a jelet információ nélkülinek tekintené! De egy matematikus számára ez pontosan azért magas információtartalmú, mert nincs benne szabályszerűség. A kép leírásához, vagy két ilyen kép megkülönböztetéséhez a szó szoros értelmében minden egyes pixelt meg kell neveznünk. Ezzel szemben a valós életből vett képeknek alacsony az információtartalma, mivel kis számú leíró elemmel továbbíthatjuk a kép tartalmát. Néhány vonás elégséges, hogy egy arc képzetét keltsük, míg egy képzett művész, viszonylag csekély számú ecsetvonással, bármilyen arcról felismerhető hasonmást tud készíteni.³

A tömörített érzékelés ötlete az, hogy a legtöbb, valós életből vett képnek pont az alacsony információtartalmát használjuk ki, és ezzel kerüljük meg a Shannon-Nyquist mintavételi tételt. Ha egyáltalán nincs információnk a helyreállítani kívánt jelről vagy képről, akkor Shannon tétele helyesen határolja be az elérhető felbontást. De ha tudjuk, hogy a kép ritkas vagy tömöríthető, akkor Shannon korlátai nem érvényesek.

Jóval azt megelőzően, hogy a „tömörített érzékelés” divatos kifejezéssé vált volna, voltak már utalások a jelenségre. Az 1970-es évek végén, földrengésekkel foglalkozó mérnökök kezdték felfedezni, hogy „az úgynevezett sarkalatos korlátok nem is sarkalatosak” - mondja Donoho. Földrengéskutatók a rétegek közötti folytonossági hiányokról visszapattanó földrengéshullámokból gyűjtenek információt a földalatti sziklaképződményekről. (A szikla állagában vagy összetételében meglevő bármilyen hirtelen változás, mint például egy olajtartalmú sziklaréteg előfordulása, a felszínre fogja tükrözni a rezgéshullámot.) Elméletileg a visszatükrözött hullámok nem tartalmaznak elegendő információt ahhoz, hogy a sziklarétegek képét

³A „képzett művész” mai változatát képtömörítő algoritmusnak hívjuk. Ilyen például a JPEG-2000 szabvány, amelyik az eredeti kép másolatát kevés számú, *wavelet*nek hívott alkotóelemből állítja elő. (Lásd a Parlez-vous Wavelets című cikket a What's happening in Mathematical Sciences 2. kötetében)

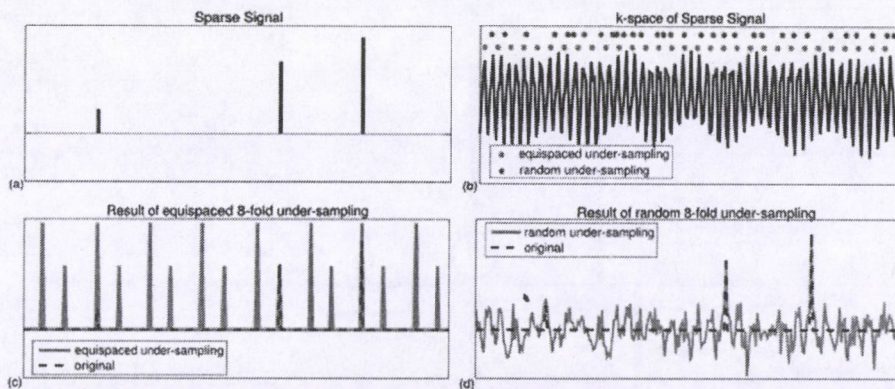


Justin Romberg. (saját szívességéből)

egyértelműen felrajzoljuk. Ennek ellenére, a földrengéskutatók képesek voltak a tőlük jogosan elvárhatónál jobb képeket nyerni. Az olajkutatás kevésbé vaktában történik azóta, hogy képesek vagyunk a föld alá „látni”. Donoho szerint a földrengéskutatók a „ritkás tüskenyaláb (sparse spike train) feltevés”-sel magyarázták szerencsésüket. A feltevés az, hogy a földalatti sziklák szerkezete elég egyszerű. A legtöbb mélységben a szikla egynemű, ezért egy bejövő földrengéshullám semmit sem lát. Szabálytalan időközönként a földrengéshullámok a szikla egyik folytonossági hiányával találják magukat szembe, és ekkor éles, tüskealakú jelet küldenek vissza küldőjüknek. Így a jel tüskék olyan ritkás sorozata, amelyben a tüskék között hosszú szünetek vannak.

Ilyen körülmények között meg lehet szabadulni Shannon tételének a béklyóitól. Könnyebb a duális helyzetet átgondolni, ami csak néhány, N hertzes frekvenciát meg nem haladó, szinuszos hullám egymásra rakásából adódó ritkás *hullámmnyaláb*ot jelent. Ha K darab frekvenciatüske van egy olyan jelben, amiben a maximális frekvencia N , akkor Shannon tétele azt mondja, hogy egyenlő közökben N mintát gyűjtsünk. De a ritkás tüskenyaláb feltevés szerint $3K$, sőt néha $2K$ minta is elégséges. A trükk abban áll, hogy nem egyenlő időközönként, hanem találomra vesszük a mintákat (lásd a 2. és 3. ábrát). Ha $K \ll N$ (és éppen ez a „ritkás” jel jelentése), akkor sokkal hatékonyabb a véletlenszerű mintavétel.

Más területeken is, mint pl. a mágneses rezonancia vizsgálat (MRI), azt találták a kutatók, hogy az adatminta „alul-mintavételezésével” is még jó eredményeket tudnak kapni. Tudományos találkozókra mindig kételkedéssel találták magukat szemben, mondja Donoho, mivel olyasmit próbáltak csinálni, ami állítólag lehetetlen volt. Visszanézve, egyfajta matematikai „tanúsítvány” hiányzott, a jóváhagyás pecsétje, ami azt garantálja, hogy a véletlenszerű mintavétel mikor működik.



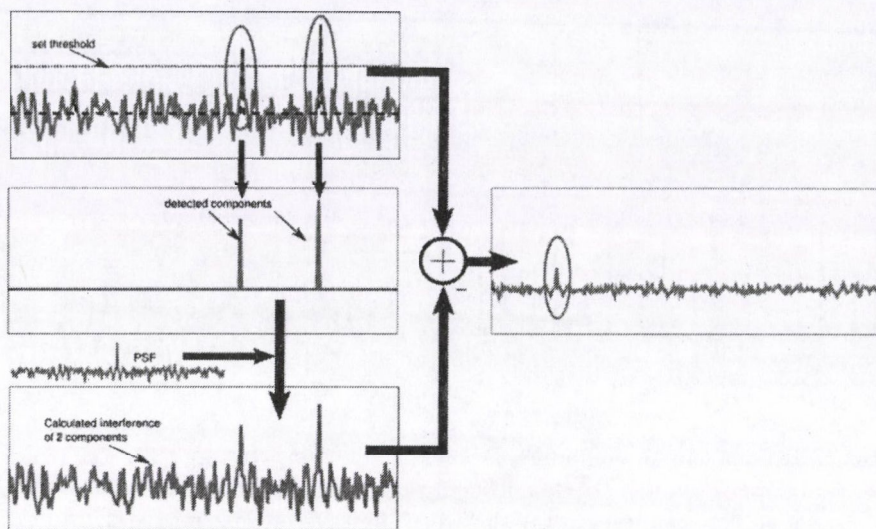
2. ábra. Ritka (sparse) hullámnyaláb rekonstrukciója. (a) Egy 3-ritka (sparse) jel frekvencia spektruma. (b) Maga a jel, két mintavételi stratégiával: egyenletes mintavétel (vörös pettyek), véletlenszerű mintavétel (kék pettyek). (c) Amikor a spektrumot az egyenletes mintavételből állítjuk vissza, súlyos "aliasing" az eredmény, mert a minták száma 8-szor kevesebb, mint a Shannon-Nyquist korlát. Lehetetlen megmondani, hogy melyik frekvencia valódi és melyik az imposztor. (d) Véletlenszerű mintavétel esetén könnyű kiszűrni a háttérből a két legnagyobb tüskét. (M. Lustig, D. Donoho, J. Santos and J. Pauly, *Compressed Sensing MRI*, Signal Processing Magazine, March 2008. szívességéből)

2. Az új bizonyítvány

Donoho egyik volt diákja, Emmanuel Candes, ugyanezzel a kételkedéssel találta magát szemben 2004-ben, amikor mágneses rezonancia vizsgálatban egy csapat radiológussal dolgozott együtt. „Fantomképes” próbafuttatásokon, azaz amikor nem egy igazi betegről készült képet futtatott, képes volt tökéletesen helyreállítani a képet alul-mintavételezett adatból. „Semmilyen különbség nem volt az eredeti és a rekonstruált kép között” - mondja Candes. „Tulajdonképpen bajba is kerültem, mert azt hitték, hogy simliskedtem.”

Ezen a ponton Candes sorsszerű dolgot tett: beszélt a 2006-os Fields-érem egyik kitüntetettjével, Terry Taoval. Úgy esett, hogy mindkét matematikus gyereke ugyanabba az óvodába járt. Miközben éppen a gyereküket adták le egyik nap, Candes mesélt Taonak a „túl jó, hogy igaz legyen” rekonstruálásokról. „Már elkezdtem keresni a magyarázatot, és némi haladást is elértem, de egy bizonyos ponton elakadtam” - mondta Candes.

„Terry matematikusként reagált” - folytatja Candes. „Azt mondta, olyan ellenpéldát fogok találni, ami megmutatja, hogy amire gondolok, az nem lehet igaz”. De furcsa dolog történt. Egyik ellenpéldáról sem látszott, hogy működne, és Tao kezdte figyelmesebben hallgatni Candes érvelését. „Egyszer csak rám nézett és



3. ábra. A 2. ábra viszonyai között a harmadik frekvencia tuskét iteratív küszöb eljárással lehet visszanyerni. Ha kezdettől tudjuk, hogy a jel 3-ritka (sparse), akkor a jelet tökéletesen vissza lehet nyerni annak ellenére, hogy 8-szoros az alulmintavétel. Röviden, ritkasság és véletlenszerű mintavétel lehetővé teszi a tökéletes (vagy majdnem tökéletes) rekonstrukciót. (M. Lustig, D. Donoho, J. Santos and J. Pauly, Compressed SensingMRI, Signal Processing Magazine, March 2008. szíveségéből)

azt mondta: Lehet, hogy igazad van”. Tao legendás gyorsaságával néhány napon belül átsegítette Candest az akadályon, és ketten elkezdtek felvázolni a tömörített érzékelés első igazán általános elméletét.

A Candes-Romberg-Tao szerkezetben egy x vektor (N valós szám sorozata) jellemez egy jelet vagy egy képet. Erről a vektorról feltesszük, hogy K -ritka (sparse), ami azt jelenti, hogy, valamilyen előírt bázisban, legfeljebb K nullától különböző együtthatót tartalmaz. (K -ról feltesszük, hogy sokkal kisebb N -nél). Például, ha a bázis tagjai $\mathbb{R}^N$ standard koordináta vektorai, akkor x szó szerint, majdnem csupa nullából áll. Pontosan ez a helyzet a ritkás túskenyaláb feltevés esetén.

A ritkás túskenyaláb feltevés azonban nem követel meg adott bázist. Például a fényképek egyáltalán nem ritkák (sparse) a standard bázisban; sok nem nulla együtthatójuk van, nevezetesen a nem fekete pixelek. A JPEG tömörítés bizonyította be, hogy egy másik bázisban viszont majdnem mindig közelítőleg ritkák (sparse), és ez a waveletek bázisa. Ha Ψ jelöli a bázisvektorok $N \times N$ -es mátrixát, akkor ebben a K -ritka (sparse) jel az lesz, amelyiket a Ψx alakban lehet írni, ahol x -nek legfeljebb K nullától különböző együtthatója van.

Az x jelből vett y mintavétel x lineáris függvénye, $y = \Phi x$, a Candes-Romberg-Tao-szerkezet szerint. A mintában levő mérések számáról, M -ről feltesszük, hogy a jel méreténél kisebb, $M \ll N$. Így Φ egy $M \times N$ -es mátrix. Elemi lineáris algebrából tudjuk, hogy végtelen sok olyan x^* vektor van, amire $y = \Phi x^*$. De ha $M \geq 2K$, akkor $y = \Phi x^*$ egyetlen más megoldása sem lesz ritka (sparse). Tehát, ha előre tudjuk, hogy x ritka (sparse), akkor az M számú mérésből elvileg pontosan rekonstruálni lehet.

Abból, hogy tudjuk, hogy csak egyetlen megoldás létezik, még nem következik, hogy képesek is vagyunk megtalálni. A probléma abban áll, hogy nem tudjuk előre, hogy x melyik K koordinátája nem nulla. A naiv megközelítés az, ha végigpróbáljuk az összes lehetőséget, amíg rálelünk a jóra, de ez reménytelenül lassú algoritmusnak bizonyul. Candes és Tao azonban egy olyan utat talált, ami nemcsak gyorsabban fut a számítógépen, de megmagyarázza azt is, hogy a véletlenszerű mintavétel miért működik sokkal jobban az egyenletes mintavételnél.

Ha képünk néhány ritkásan elhelyezkedő pontból vagy néhány éles vonalból áll, akkor a *legrosszabb* módszer az, ha az egyes pixelek befogásával veszünk mintát (a közönséges fényképezőgép pont így működik). A *legjobb* módszer az, ha a képet szélesen kiterített zajfüggvényekkel vetjük össze. A „20 kérdés” nevű játékkal⁴ tudjuk ezt érzékeltetni. Ha 1-től N -ig kell egy számot kitalálni, a legrosszabb amit tehetünk, ha találgatjuk az egyes számokat (ez felel meg az egyes pixelek megmérésének). Átlagban ez $N/2$ találgatást jelent. Ezzel szemben, ha olyan kérdéseket teszünk fel, hogy „a szám kisebb $N/2$ -nél?”, azután „a szám kisebb $N/4$ -nél?” és így tovább, akkor az elrejtett számot legfeljebb $\log_2 N$ kérdéssel megtalálhatjuk. Ha N nagy szám, akkor ez hatalmas gyorsulás.

Figyeljük meg, hogy a „20 kérdés” stratégia adaptív: az előző válaszok fényében szabad változtatnunk kérdéseinken. Candesnek és Taonak nem adaptív eljárást kellett kifejlesztenie ahhoz, hogy a gyakorlatban is használható legyen, de úgy, hogy közben ugyanazt a garantált teljesítményt nyújtja, mint az előbb leírt adaptív stratégia. Más szóval, előre ki kellett találniuk, hogy mik lennének az x -ről legtöbb információt nyújtó kérdések. Az új elmélet egyik nagy meglepetése, hogy ezt meg is lehet tenni. Megközelítésükben az ötlet az ℓ_1 minimalizálás.

Egy vektor ℓ_0 normája egyszerűen a benne levő, nullától különböző bejegyzések száma, amit kicsit fesztelenül így lehet írni:

$$\| (x_1, x_2, \dots, x_n) \|_0 = \sum |x_j|^0.$$

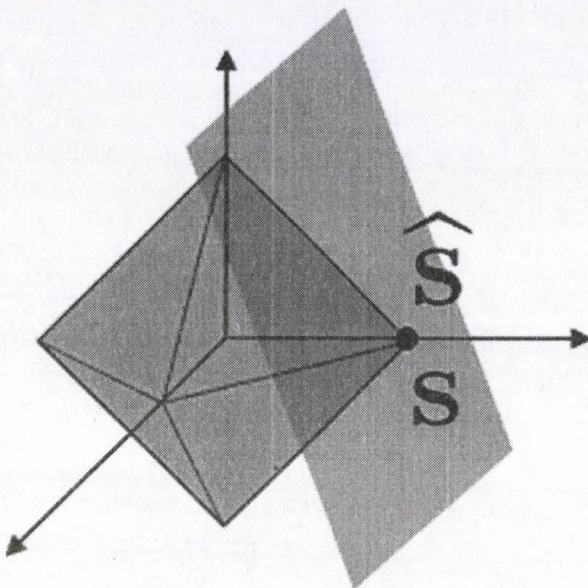
(Ebben a formulában, megállapodás szerint, $0^0 = 0$.) Az ℓ_1 normát úgy nyerjük, hogy a 0-akat 1-esekre cseréljük az előző egyenletben:

$$\| (x_1, x_2, \dots, x_n) \|_1 = \sum |x_j|.$$

Ezen a nyelven az $y = \Phi x^*$ egyenlet egyetlen legkisebb ℓ_0 normájú megoldása éppen x . De Tao és Candes megmutatta, hogy sok esetben ez az x az egyetlen

⁴A barkochba játék egy amerikai, televíziós változata, en.wikipedia.org/wiki/Twenty_Questions

legkisebb ℓ_1 normájú megoldás is. Ez a felismerés döntőnek bizonyult, mert az ℓ_1 minimalizálás lineáris programozási feladat, amit ismert, hatékony számítógépes algoritmusokkal meg lehet oldani. (Lásd a Smooth(ed) Moves című cikket a What's happening in Mathematical Sciences 6. kötetében)



4. ábra. Az S ritka (sparse) jel véletlenszerű mérése ezt a mérést produkáló lehetséges jelek (zöld) alterét generálja. A zöld altérben belül a legkisebb ℓ_1 normájú vektor, $S^\wedge$, általában S -sel egyenlő. (R. G. Baraniuk, Compressive Sensing [Lecture Notes], Signal Processing Magazine, July 2007. szívésségéből)

A 4. ábra illusztrálja, hogy miért gyakori az, hogy az ℓ_1 -et minimalizáló ugyanaz, mint az ℓ_0 -t minimalizáló. A 3 dimenziós térben az ℓ_1 norma szerinti egységvektorok egy oktaédert alkotnak. Mivel sok nulla koordinátája van, gondoljuk el, hogy a ritka (sparse) x vektor benne fekszik egy koordináta tengelyben, ezért az oktaéder egyik csúcsában végződik. Azon x^* vektorok halmaza, amelyekre $y = \Phi x^*$ teljesül, egy, az x ponton átmenő sík. A legtöbb az x ponton átmenő sík az oktaédert *csak* ebben az egyetlen x pontban metszi, másképpen szólva, x a síkon az egyetlen minimális normával rendelkező pont. Tehát, ha egyszerűen a Φ mérést véletlenszerűen tesszük meg, jó esélyünk van rá, hogy x -et egyértelműen rekonstruálni tudjuk.

Sajnos Φ véletlenszerű kiválasztása nem mindig működik. Esetleg pechesek leszünk, és pont egy olyan x -en átmenő síkot választunk, ami az oktaéder belsejébe is belemetsz. Ha ez történik, akkor az ℓ_1 minimalizáló nem lesz ugyanaz, mint az ℓ_0 minimalizáló. Az algoritmus egy hibás x^* jelet fog eredményezni. De a 4. ábrán látható kép valamennyire félrevezető, mivel a képek vektorainak dimen-

ziója általában ezres vagy milliós nagyságrendű. A millió-dimenziós térben az oktaéder analógját keresztpolitópnak hívják; és a keresztpolitóp a millió-dimenziós térben nagyon, nagyon, nagyon hegyes. Egy a csúcson átmenő, de egyébként véletlenszerű sík *majdnem biztos*, hogy a keresztpolitóp belsejét elkerüli. A „nagy dimenziós geometria e csodájának” köszönhetően, ahogy ezt Candes nevezi, az ℓ_1 minimalizáló majdnem biztosan a korrekt jel lesz.

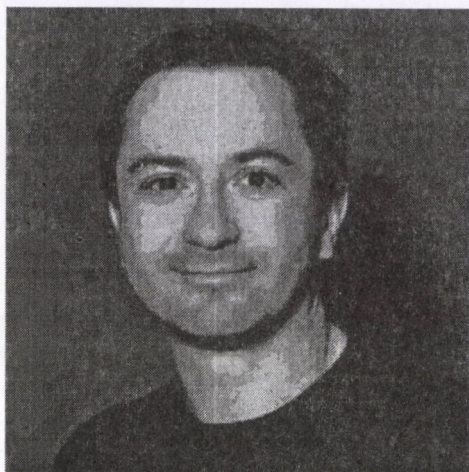
Összefoglalva, a tömörített érzékelés elmélete a következőket mondja:

- „Sok Φ $M \times N$ -es mátrixra az $y = \Phi x^*$ egyenlet egyetlen K -ritka (sparse) megoldását *pontosan* vissza lehet nyerni.”
- N -nek K -nál sokkal nagyobbnak kell lennie, azonban M -nek, a mérések számának csak egy kicsit kell K -nál nagyobbnak lennie. Pontosabban, M -nek durván $K \log(N/K)$ -nak kell lennie. Figyeljük meg, hogy ez N -től logaritmikusan függ, így a „20 kérdés”-nél véghezvitt gyorsulást most is elértük.
- „A K -ritka (sparse) megoldást ℓ_1 minimalizálással találjuk meg, amiről a Φ mérési mátrixra vonatkozó bizonyos feltételek teljesülése esetén belátható, hogy az ℓ_0 minimalizálással ekvivalens.”
- A Φ véletlen mátrixok majdnem mindig kielégítik ezeket a feltételeket.

Az egész történet lényegében változatlanul megmarad, ha a jel nem a standard koordináta vektorok bázisára nézve, hanem egy másik bázisra, mondjuk a waveletekére nézve ritka (sparse). Az egyetlen dolog, ami változtatásra szorul az, hogy az $y = \Phi x^*$ megszorítást az $y = \Phi \Psi x^*$ megszorítással helyettesítjük. Ebben az összefüggésben a Φ mérési mátrix véletlenszerűsége kettős célt szolgál. Először, azt a legegyszerűbb feltételhalmazt biztosítja, amikor az ℓ_1 minimalizálás bizonyíthatóan ekvivalens az ℓ_0 minimalizálással. Másodszor pedig, az előbbtől függetlenül biztosítja, hogy a mérési vektorok halmaza (Φ sorai) annyira különböznek a kép bázisától (Ψ oszlopaitól), amennyire ez csak lehetséges. Ha a kép bázis tüskékből áll, akkor a mérések bázisának kiterjedt véletlenszerű zajból kellene állnia. Ha a kép bázis waveletekből áll, akkor a mérések bázisa „noiselet”-eknek hívható kiegészítő típusú jelből kell állnia.

„Cikkünk valami tényleg váratlanra mutatott rá” - mondja Candes. „Arra nevezetesen, hogy rendkívül hatásos érzékelő mechanizmust ad a véletlenszerűség használata. Ez az első számú állítás. A második számú állítás pedig az, hogy mindez kiállja a szigorú elemzés próbáját.”

„Ami a matematikusoknak a cikkből tetszett, az a mód volt, ahogy az analízist és a valószínűségi számítást egybeolvasztotta. Területemen, az analízisben sokan úgy gondolták, hogy a valószínűségi számítás nem hasznos vagy nem méltó a figyelmükre. Intellektuális szinten változtatta meg a cikk gondolkodásukat, és arra készítette őket, hogy foglalkozzanak ezzel a területtel is.”



Richard Baraniuk. (saját szívességéből)

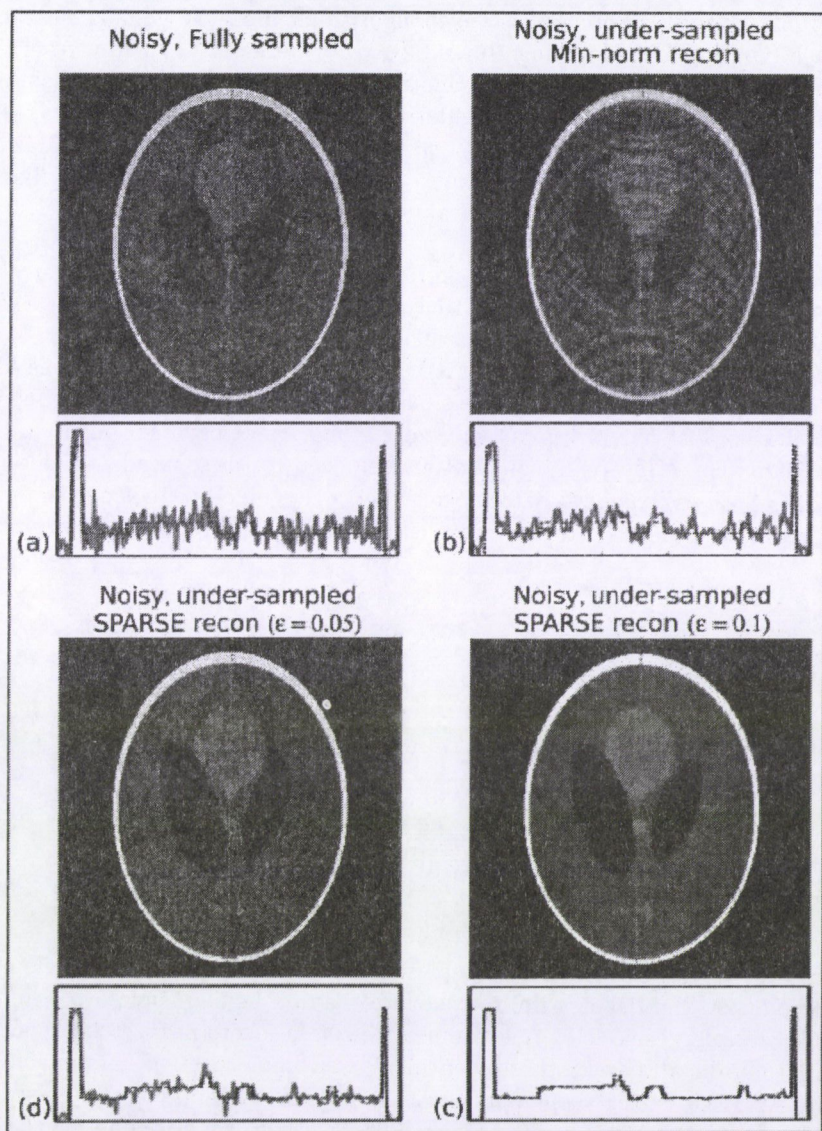
3. Legújabb fejlemények

Tao és Candes preprintje 2004-ben jelent meg, mint ahogy Donoho hasonló eredményeket közlő cikke is. Mire Tao és Candes cikke 2006-ban ténylegesen megjelent, már több, mint 100-szor idézték. Azóta mind elméleti, mind gyakorlati téren sok előrelépés történt.

Az eredeti cikkben megválaszolatlanul hagyott kérdés az volt, hogy mennyire marad jó a tömörített érzékelés, ha a mérés valamilyen véletlenszerű hibát tartalmaz (valódi eszközöknél ez elkerülhetetlen probléma), vagy ha a ritka (sparse) jelről szóló feltételezés nem teljesen igaz. Fényképezésnél például szó szerint nem igaz. Élethűbb az a feltételezés, hogy a jel tömöríthető, ami azt jelenti, hogy a jel néhány együtthatója tartalmazza az információ túlnyomó részét. A többi együttható nem teljesen nulla, de kicsi. Ilyen körülmények között még az ℓ_0 minimalizálás se illik pontosan össze a jellel, így remény sincs rá, hogy az ℓ_1 minimalizálás teljesen korrekt legyen.

Candes, Romberg és Tao megmutatta 2005-ben, hogy a tömörített érzékelés jól működik még zajos mérések és tömöríthető, bár nem ritka (sparse) jelek esetében is. A helyreállított jel hibája nem lesz sokkal nagyobb a mérések hibájánál, és az ℓ_1 minimalizálás nem lesz sokkal nagyobb annál a büntetésnél, amit az ℓ_0 minimalizálás már ránc rött. Vagyis az ℓ_1 minimalizálás hűen adja vissza a legfontosabb információkat, a jel legnagyobb összetevőit. Az 5. ábra szimulált képhez adott zaj esetén ad példát a tömörített érzékelés teljesítményére.

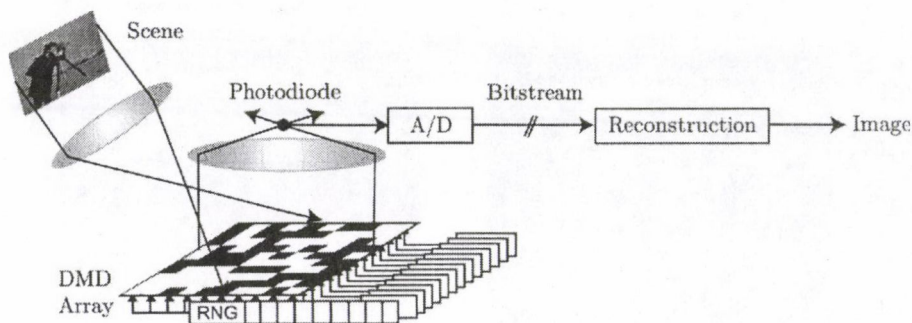
A matematikusok olyan, a szokványos lineáris programozási eljárásoknál gyorsabb, új algoritmusokon dolgoznak, amelyek megoldják az ℓ_1 minimalizálás problémát. x legnagyobb K együtthatóját nem mind egyszerre találják meg, hanem iteratívan: először a legnagyobb nullától különböző együtthatót, azután a második



5. ábra. Tömörített érzékelés zajos adatokkal. (a) Kép hozzáadott zajjal. (b) Az alul-mintavételezett, majd a Shannon-Nyquist megközelítéssel helyreállított kép. (d) Ugyanaz a kép, véletlenszerűen alul-mintavételezve, majd egy "túlságosan optimista" zaj modellel helyreállítva. Habár nincsenek műleletek, a zaj egy részét félreértette és valós módosulásként vette az eljárás. (c) Ugyanaz a kép, egy toleránsabb zaj modellel rendelkező véletlenszerű mintából helyreállítva. Az eljárás elnyomta a zajt és nincsenek műleletek sem. (Michael Lustig szívességből)

legnagyobbat, és így tovább. Az első ilyen algoritmus, amit ortogonális illeszkedési keresésnek (Orthogonal Matching Pursuit (OMP)) neveztek el, nem érte el az ℓ_1 minimalizálás által biztosított pontosságot. De mára már az OMP változatainak színes csokra létezik, mint például regularizált OMP (ROMP), vagy a szakaszonkénti, „stagewise” OMP (StOMP), amik sikeresen ötvözik az ℓ_1 minimalizálás pontosságát és az OMP gyorsaságát. Ezeknek az algoritmusoknak az az előnyük az ℓ_1 minimalizálás „nagy dimenziós csodájával” szemben, hogy intuitívabbak. A 3. ábra mutat erre példát.

Eközben több különböző terület kutatója puhatojja a tömörített érzékelés gyakorlati alkalmazásait. Baraniuk és Kelly 2006-ban épített egyetlen pixeles fényképezőgépe egy sereg baktérium méretű tükröt használ, hogy a bejövő fényből véletlenszerű mintát vegyen. (Lásd a 6. ábrát.) Mindegyik tükröt kétféle állás egyikébe lehet dönteni: vagy az egyetlen érzékelő felé tükrözi a fényt, vagy elfelé. Tehát az érzékelő által kapott fény sok pixel súlyozott átlaga, amit egyetlen pixelben gyűjtöttünk össze. Ha $K \log(N/K)$ pillanatfelvételt készítettünk, minden alkalommal különböző, véletlenszerűen választott pixelekkkel, az egypixeles fényképezőgép képes volt felismerhető képet készíteni, N pixellel összemérhető felbontásban.



6. ábra. Az "egypixeles fényképezőgép" sematikus diagramja. Az érzékelő egyetlen fotódióda. A "DMD" egy mikro-tükrös rács, ami a bejövő fénycsóva egy részét visszatükrözi az érzékelő felé. A kép másik részét (a fekete négyzeteket) elfelé téríti. A fotódióda által végzett minden mérés sok pixel véletlen kombinációjából jön létre. Az "Egy is elég" képen 1600 véletlen mérés elegendő volt, hogy egy 4096 pixeles fényképezőgép által készített képpel összevethető képet csináljon. (Richard Baraniuk szívességéből)

Baraniuk és Kelly csapata most a „hiperszínképes fényképezőgépen” dolgozik, amely a kép minden egyes pontjában a teljes színeképet rekonstruálná. „Egy hagyományos digitális képnek vannak vörös, kék és zöld pixellei.” - mondja Baraniuk. „Ez nagyszerű arra, hogy olyan fényképet csináljunk, ami becsapja az emberi szemet, de ez nem fogja be a különböző anyagok által kiadott hullámhosszok lényegét. Igazából ezernyi színből álló spektrumot szeretnénk kapni pusztán három szín

helyett. Ez lehetővé tenné, hogy meg tudjuk különböztetni egy kocsi zöld festékét és egy bokor zöld levelét.” Mivel a milliónyi pixel mindegyike ezernyi színű lehet, adattömörítés komoly üggyé válik.

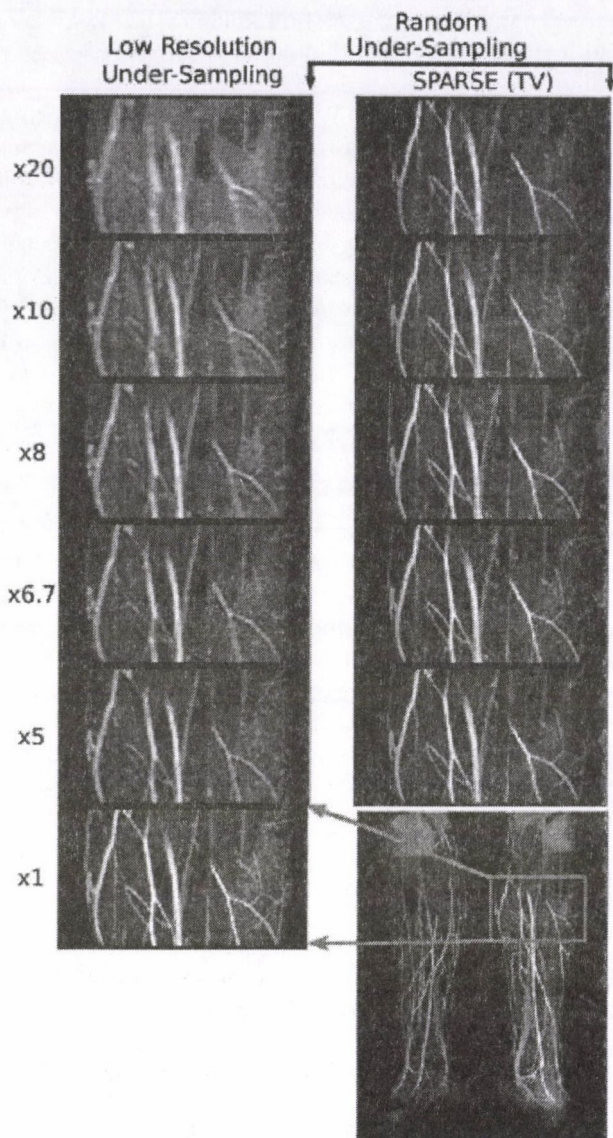
Baraniuk volt diákjával, Michael Wakinnal, aki most a Michigan Egyetemen van, szintén dolgozott a tárgyészlelési problémán. Sok alkalmazásnál nem egy tényleges fénykép elkészítése a fontos, hanem az, hogy gyorsan felismerjük, hogy mi van benne. Például, egy biztonsági rendszernek esetleg egy arcot vagy egy járművet kell azonosítania. Megtaníthatjuk, hogy felismerje egy Corolla és egy Porsche közti különbséget - példálózik Baraniuk. A számítógépben Corolla és Porsche képek vannak tárolva, de a fényképezőgép előtti jármű lehet, hogy úgy van elfordulva, hogy egyik képpel sem illik pontosan össze. Ebben az alkalmazásban a képvektornak más jellegű ritka (sparse) szerkezete van. Nem egy K -dimenziós síkon fekszik, hanem az N -dimenziós térben levő görbült K -dimenziós sokaságon. (Esetünkben K 3-mal egyenlő.) Wakin megmutatta, hogy ebben a környezetben $K \log N$ mérés még mindig elégséges.

A tömörített érzékelés néhány alkalmazása teljesen kívül esik a képfeldolgozás birodalmán. Az egyik ilyen példa az „analógról digitálisra konvertálás”, ami a drótnélküli, mobil kommunikáció alapvető jellegzetessége. Például a CDMA mobiltelefon szabvány bevesz egy hangüzenetet, ami 4096 hertzig terjedő hangfrekvenciákat tartalmaz, és szétteríti a rádióspektrumban, ami százezernyi hertzig terjed. A jel ritka (sparse), mivel csak annak a 4096 hertznek a belsejébe préselt információt tartalmaz. Tehát egy tömörített érzékelést végrehajtó vevőnek gyorsabban kell visszanyernie a jelet, mint egy Shannon tételén alapuló vevőnek.

Moore szabálya szerint a digitális fotográfiában kétévenként duplázzuk meg az egy áramkörti lapkán lévő jelfogók számát. De a digitális átalakítás világában, Baraniuk szerint „az egyenértékű érdemleges szám 6-8 évenként duplázódik meg”. Tehát ne várjunk évtizedeket hardver megoldásra, hanem oldjuk meg a problémát a tömörített érzékelésen alapuló szoftverrel.

Végül, a tömörített érzékelésnek orvosi alkalmazásai is vannak. Ez nem meglepő, hiszen egy, a mágneses rezonancia vizsgálatban talált probléma közvetlenül ihlette elméletének létrejöttét. Az MRI készülékek használata hagyományosan mozdatlan struktúrák képének rövid idő alatti létrehozására korlátozódott, és a beteget utasították, hogy tartsák vissza a lélegzetüket. De mostanában, a képet időben és térben ritka (sparse) jelként kezelő MRI készülékek elkezdtek leküzdeni ezeket a korlátokat, például egy dobogó szív képét állítják elő. A 7. ábra mutatja, hogy egy ritka (sparse) rekonstrukciós algoritmus mennyire képes éles képet nyújtani a beteg lábában lévő artériákról, pedig 20-szor kevesebb adatot használ, mint egy hagyományos angiogram.

Egy akadályt még le kell küzdenie a tömörített érzékelésnek. Ez azt jelenti, hogy hogyan fejlesszen ki gyakorlatban használható „következtelen” érzékelőket. A tömörített érzékelésben egyetlen mérés a bejövő tömöríthető jel, és a véletlenszerű, zajos teszt jel skaláris szorzata. Baraniuk egypixeles fényképezőgépe úgy hajtja végre a skaláris szorzást, hogy a fénycsóva egy részét a tükrök az érzékelő felé, más részét elfelé továbbítják. Ha a következtelen méréseket lehetővé tevő



7. ábra. Egy angiogram. Felülről lefelé, az angiogramot fokozatosan egyre nagyobb tényezőkkel fokozatosan alul-mintavételezzük. A Shannon-Nyquist mintavételezési stratégiát használva, a kép minősége úgy csökken, ahogy az alul-mintavételezés nő. A tömörített érzékelést használva, a kép még 20-szoros alul-mintavételezés esetén is nagyon éles körvonalú marad. Az itt és az 5. ábrában használt megközelítés nem ℓ_1 minimalizálás, hanem a térbeli gradiens ℓ_1 minimalizálása. (Michael Lustig szívességéből)

hardver drágább, mint a felváltani kívánt érzékelők sokasága, akkor a tömörített érzékelés melletti gazdasági érvek elenyésznek valós alkalmazásokban.

Miután annyi kérdés és annyi lehetőség van, pillanatnyilag lehetetlen megmondani, hogy melyik lesz a tömörített érzékelés legsikeresebb alkalmazása. Azonban egy dolog világos: a mérnökök végre kilépnek a Shannon tételét jelentő dobozból.