Compressed and Sparse Sensing

Yucheng Wu

2024/5/24

目录

- 压缩感知（compressed sensing, CS）
  - 内容来源：《形象易懂讲解算法II——压缩感知》（https://zhuanlan.zhihu.com/p/22445302）

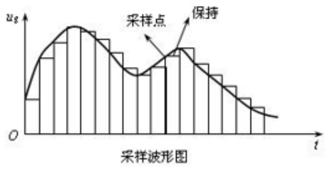
- 稀疏感知（Sparse sensing）
  - Sparse Mobile Crowdsensing: Challenges and Opportunities (IEEE Communications Magazine, 2016)

一、什么是压缩感知?

- compressed sensing又称compressed sampling，是一个针对信号采样的技术，它通过一些手段，实现了“压缩的采样”，**在采样过程中完成了数据压缩的过程。**

- 信号采样：

模拟信号采样



问题 用多大的采样频率？

奈奎斯特采样定理

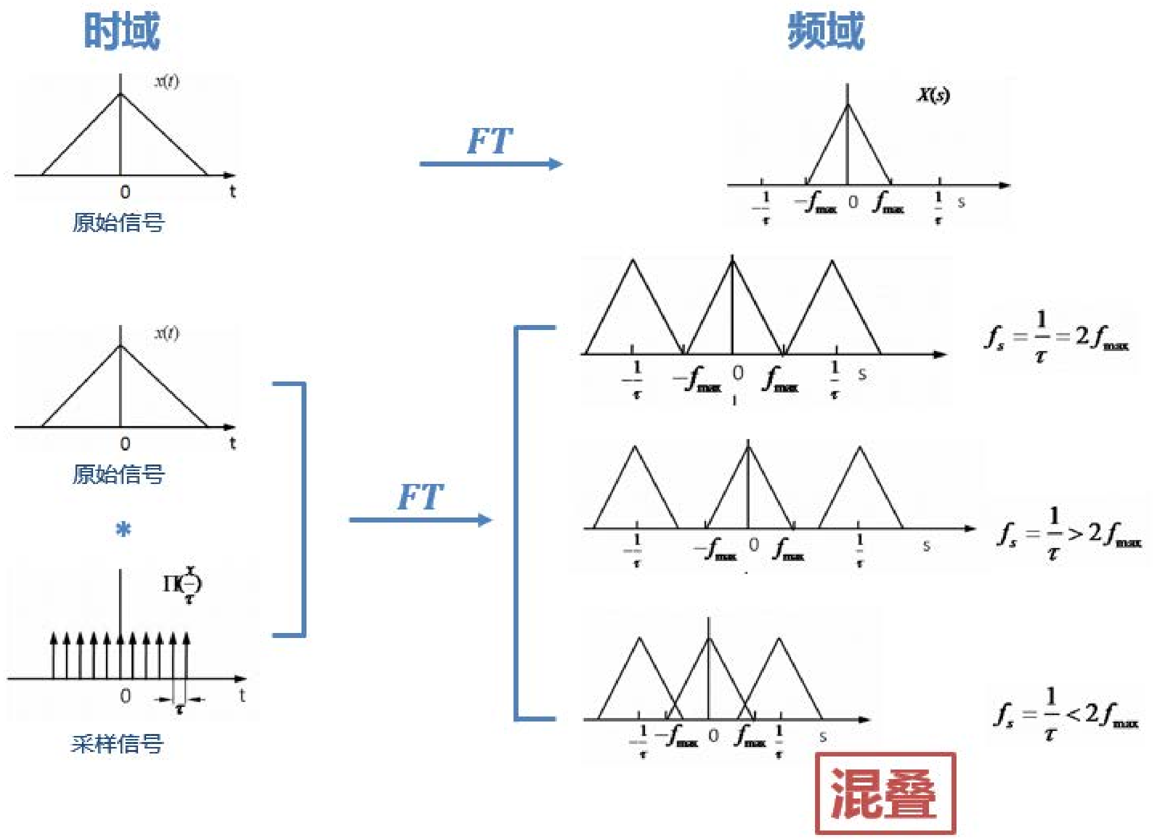
奈奎斯特，H.

1889-1976

要想让采样之后的数字信号完整保留原始信号中的信息，采样频率必须**大于信号中最高频率的2倍**！

奈奎斯特采样定律

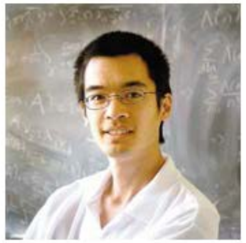
- 为什么是两倍：时域以τ为间隔进行采样，频域会以1/τ为周期发生周期延拓。那么如果采样频率低于两倍的信号最高频率，信号在频域频谱搬移后就会发生**混叠**。



一定要如此吗？

2004年，几位大牛证明，如果信号是稀疏的，那么它可以由远低于采样定理要求的采样点重建恢复，并于2007年正式提出了“压缩感知”（Compressed Sensing）这个概念。

压缩感知



陶哲轩

Emmanuel Candes

David Donoho

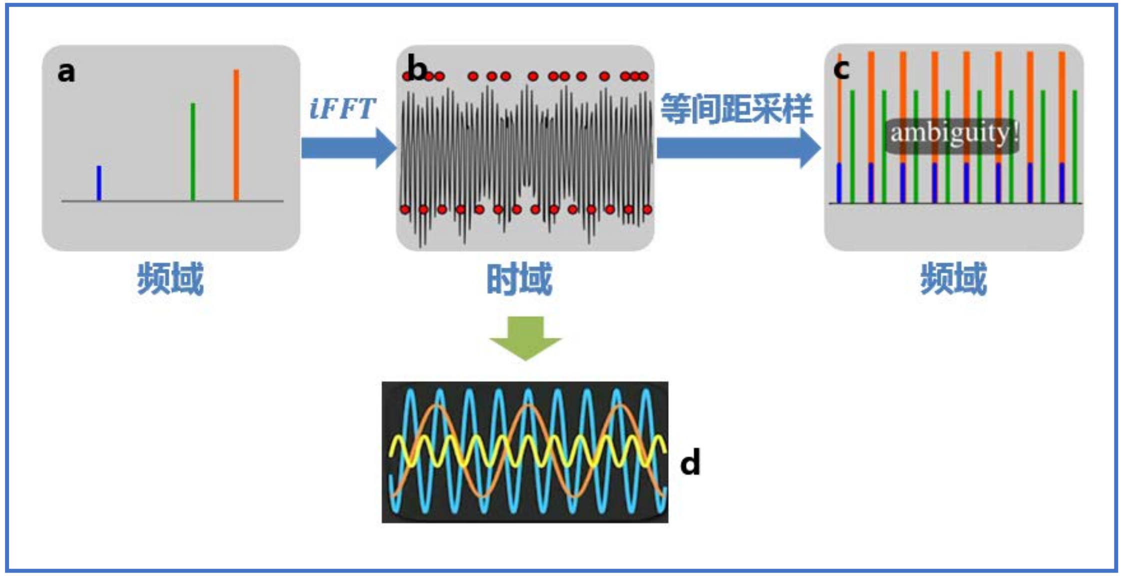
压缩感知

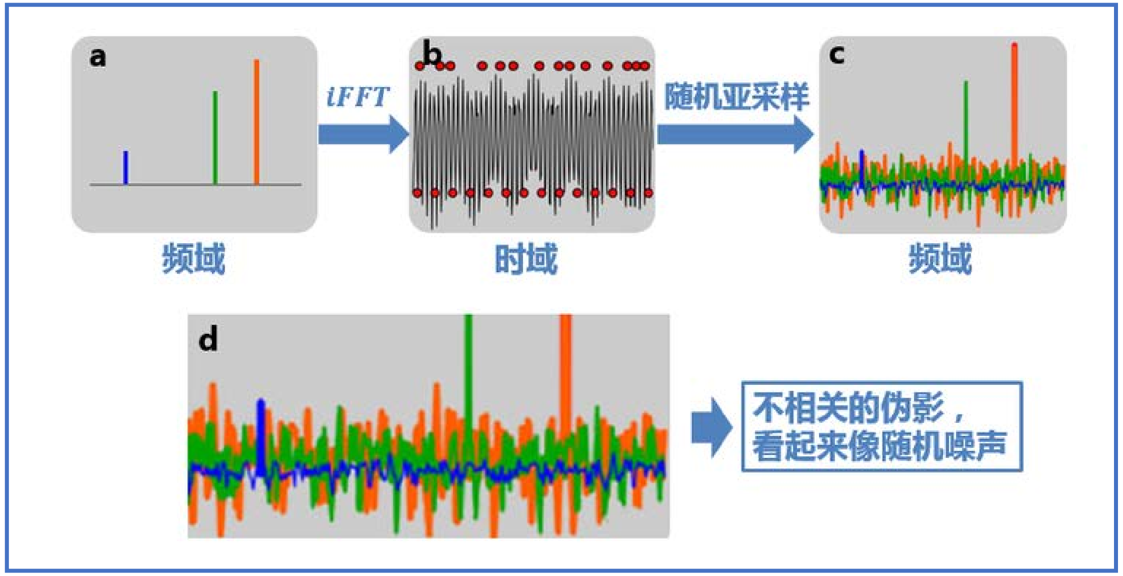
“采样频率须大于信号中最高频率的2倍”——Nyquist

采样频率 $\xrightarrow{\text{意味着}}$ 等间距采样

- 等间距采样，频域将以1/τ为周期延拓，采样频率低必然引起混叠。
- 如果是不等间距采样呢？
- 如果是随机采样呢？

随机的亚采样给了我们恢复原信号的可能

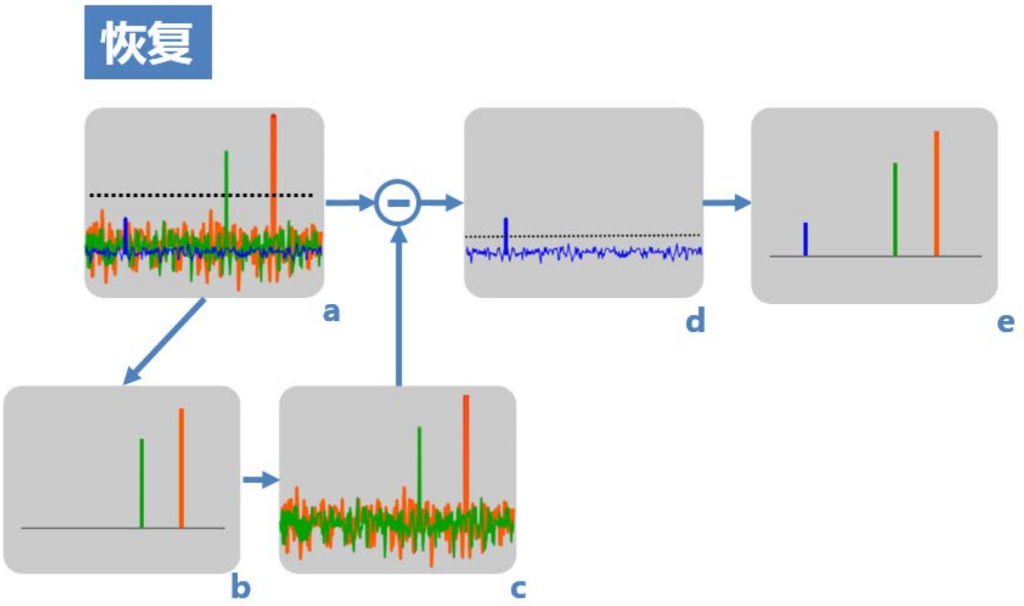




- 如图b、d为三个余弦函数信号叠加构成的信号，在频域的分布只有三条线（图a）。如果对其进行8倍于全采样的等间距亚采样（图b下方的红点），则频域信号周期延拓后，就会发生混叠（图c），无法从结果中复原出原信号。
- 而如果采用**随机**亚采样（图b上方的红点），那么这时候频域就不再是以固定周期进行延拓了，而是会产生大量不相关（incoherent）的干扰值。如图c，**最大的几个峰值还依稀可见**，只是一定程度上被干扰值覆盖。这些干扰值看上去非常像随机噪声，但实际上是由于三个原始信号的非零值发生能量泄露导致的（不同颜色的干扰值表示它们分别是由于对应颜色的原始信号的非零值泄露导致的）
- P.S：为什么随机亚采样会有这样的效果？这可以理解成随机采样使得频谱不再是整齐地搬移，而是一小部分一小部分胡乱地搬移，频率泄露均匀地分布在整个频域，因而泄漏值都比较小，从而有了恢复的可能。

信号该如何恢复?

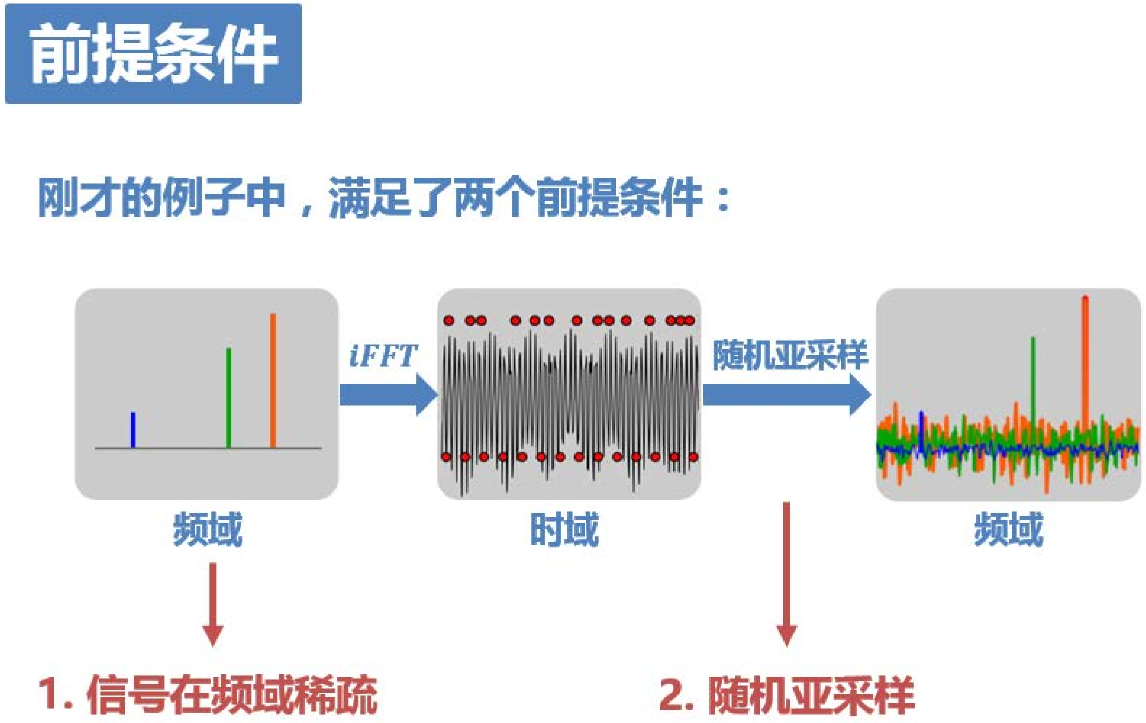
- 一种典型的算法（匹配追踪）：
  1. 由于原信号的频率非零值在亚采样后的频域中依然保留较大的值，其中较大的两个可以通过设置阈值，检测出来（图a）。
  2. 然后，假设信号只存在这两个非零值（图b），则可以计算出由这两个非零值引起的干扰（图c）。
  3. 用a减去c，即可得到仅由蓝色非零值和由它导致的干扰值(图d)，再设置阈值即可检测出它，得到最终复原频域（图e）
  4. 如果原信号频域中有更多的非零值，则可通过迭代将其一一解出。



- 以上就是压缩感知理论的核心思想——以比奈奎斯特采样频率要求的采样密度更稀疏的密度对信号进行随机亚采样，由于频谱是**均匀泄露**的，而不是**整体延拓**的，因此可以通过特别的追踪方法将原信号恢复。

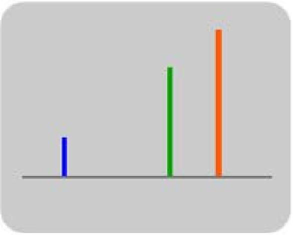
二、压缩感知的前提条件

- 1. 这个信号在频域**只有3个非零值**，所以可以较轻松地恢复出它们。
- 2. 采用了**随机亚采样机制**，因而使频率泄露均匀地分布在整个频域。
- 这两点对应了CS的两个前提条件——**稀疏性**（**sparsity**）、**不相关性**（**incoherence**）。



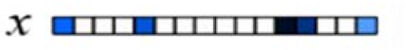
前提条件1：稀疏性

信号需要在某一个变换域具有稀疏性



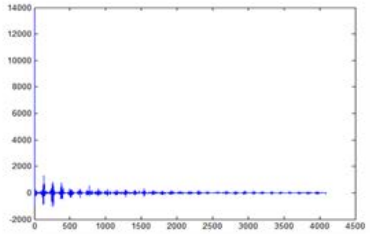
信号在频域是稀疏的

如果信号在某个域中非零点远远小于信号总点数，则信号在该域中是稀疏的。



典型的一维稀疏信号，只有少量非零项

对于压缩感知：



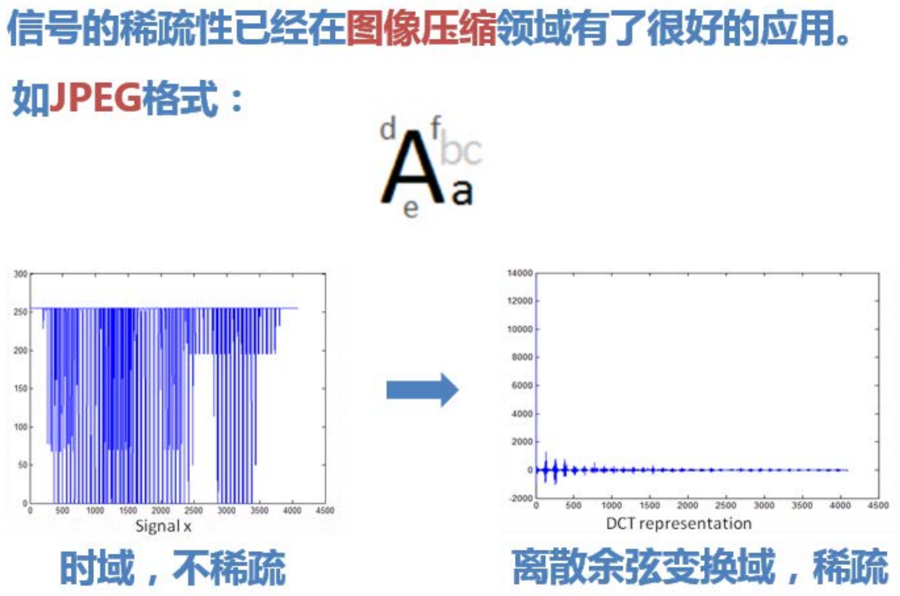
信号只需要近似满足稀疏性，即为可压缩信号。

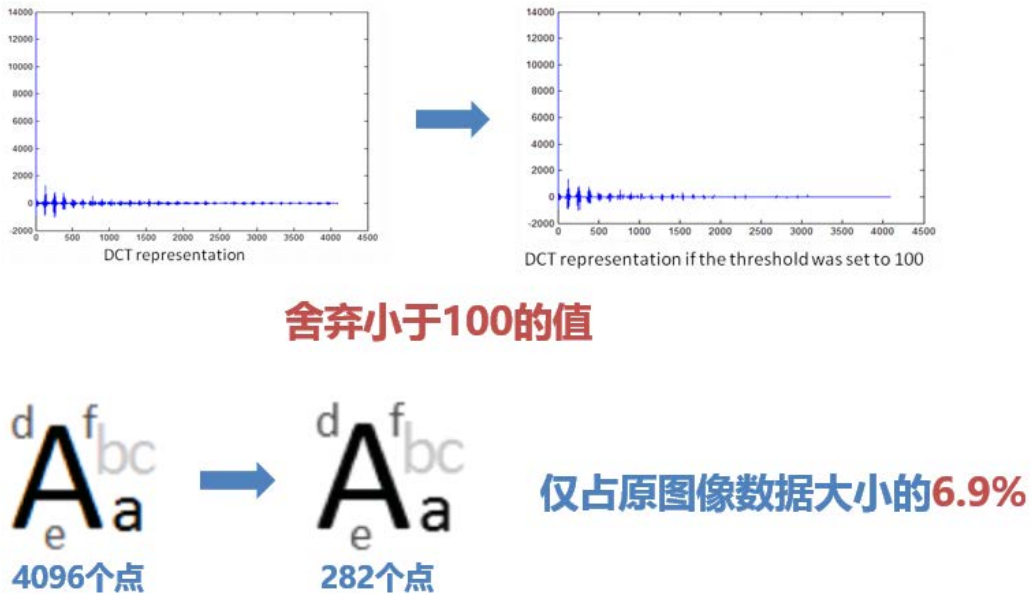
对于CS，只要它在某一个变换域满足近似稀疏特性即可，我们称之为稀疏域，重建将在稀疏域进行。

Ψ可以是频域、小波变换、离散余弦变换等

补充知识：稀疏性

- 利用信号的稀疏性，可以对信号进行压缩。如图像压缩领域的JPEG格式，就是将图像变换到离散余弦域，得到近似稀疏矩阵，只保留较大的值，从而实现压缩。

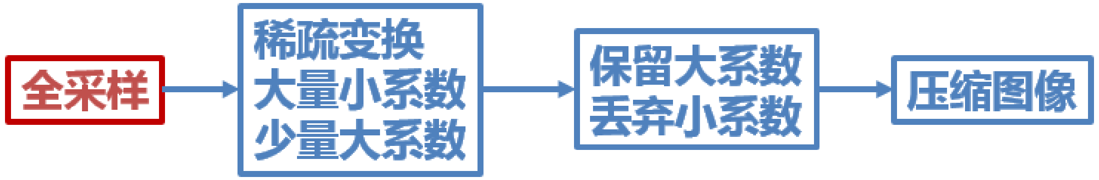




图像压缩 vs 压缩感知

- 图像压缩：先进行了**全采样**，然后再变换域丢弃小系数，完成压缩
- 压缩感知：直接进行了**亚采样**，然后再用算法消除亚采样导致的伪影。可以说，压缩感知**直接在采样时就完成了压缩**。

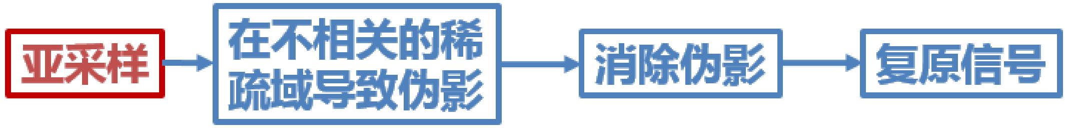
图像压缩



压缩感知的原动力问题：

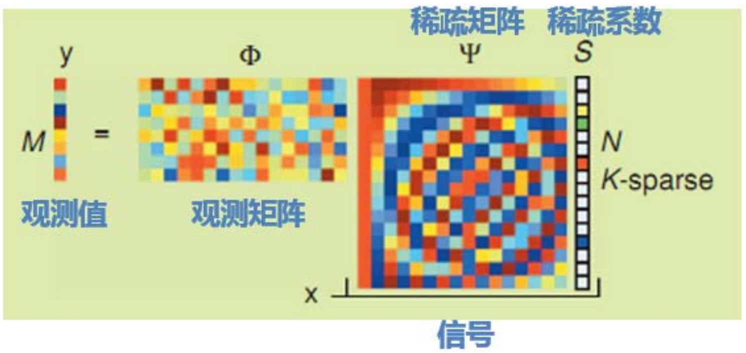
既然全采样了还要再丢弃，我们为什么不能直接少采样一些点？

压缩感知



数学表达

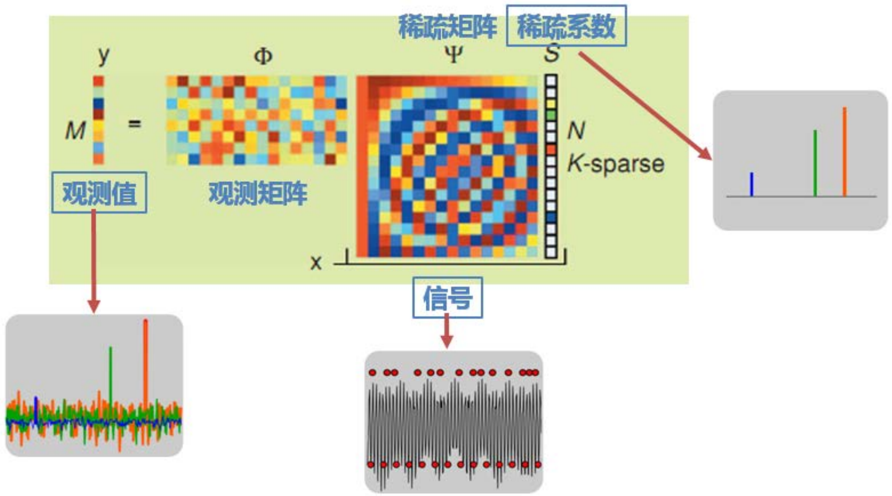
- x是为长度N的一维信号，也就是原信号，稀疏度为k。此刻它是未知的。
- Φ为观测矩阵，对应着亚采样这一过程。它将高维信号x投影到低维空间，是已知的。
- y=Φx为长度M的一维测量值，也就是亚采样后的结果。显然它也是已知的。
- 因此，**压缩感知问题就是在已知测量值y和测量矩阵Φ的基础上，求解欠定方程组y=Φx得到原信号x。**
- 令x=Ψs，Ψ为稀疏基矩阵，s为稀疏系数。
- 于是最终方程就变成了：**y=ΦΨs。已知y、Φ、Ψ，求解s。**



$$y = \Phi\Psi s$$

x ：信号
Φ ：观测矩阵
y ：观测值

- 观测矩阵Φ将高维信号x投影到低维空间。
- 对x在Ψ稀疏基上进行稀疏表示，x=Ψs，Ψ为稀疏基矩阵，s为稀疏系数。

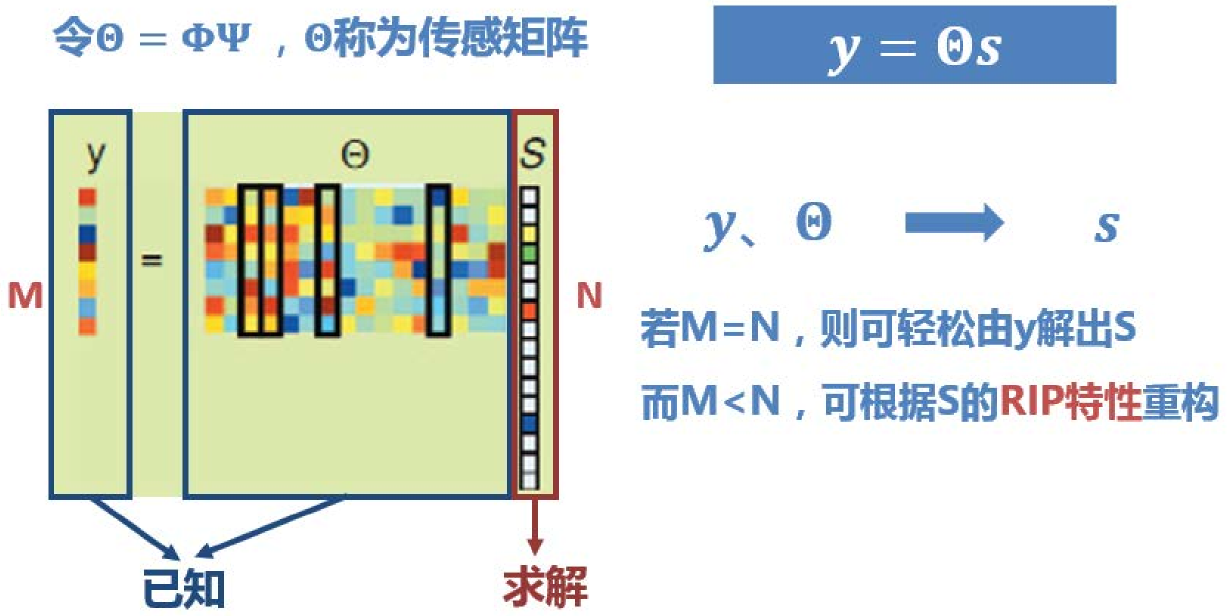


x就是三个正弦信号叠加在一起的原信号；稀疏矩阵Ψ就是傅里叶变换，将信号变换到频域S；而观测矩阵Φ就对应了我们采用的随机亚采样方式；y就是最终的采样结果。

数学表达

- 把ΦΨ合并成一个矩阵，称之为传感矩阵。即令Θ=ΦΨ，则**y=ΘS**。
- **问题即为，已知y和Θ，求解S。**
- 求解出S后，由x=Ψs即可得到恢复出的原信号x。

- 然而在正常情况下，方程的个数远小于未知数的个数，方程是没有确定解的，无法重构信号。但是，由于信号是K稀疏，如果上式中的Φ满足有限等距性质(RIP)，则K个系数就能够从M个测量值准确重构（得到一个最优解）。



前提条件2：不相关

之前已提到，采用**随机**亚采样才能实现信号的恢复。

陶哲轩和Candès于2005年给出了更为准确的要求：观测矩阵Φ应满足约束等距性条件(Restricted Isometry Property，简称**RIP**)：

即对于任意和常数，有:

$$(1-\delta_k)\|c\|_2^2 \le \|\phi c\|_2^2 \le (1+\delta_k)\|c\|_2^2$$

Baraniuk证明：

RIP的等价条件是观测矩阵和稀疏表示基**不相关**（**incoherent**）

$$y = \Phi\Psi s$$

$$\Phi \xleftrightarrow{\text{不相关}} \Psi$$

相关性的定义：

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

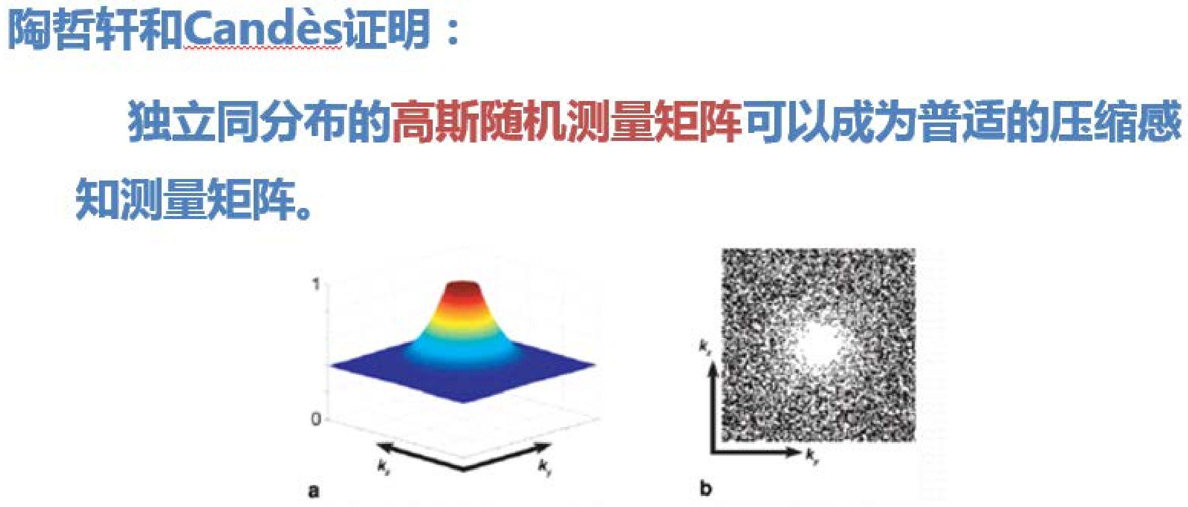
μ的范围： $\mu(\Phi,\Psi) \in [1, \sqrt{n}]$

μ越小，Φ和Ψ越不相关

前提条件2：不相关

- 那怎样找到不相关的观测矩阵呢？陶哲轩和Candès又证明：**独立同分布的高斯随机测量矩阵可以成为普适的压缩感知测量矩阵。**

- 于是满足高斯分布的随机测量矩阵就成了CS最常用的观测矩阵。
- 对于二维信号，往往就采用如图所示的采样矩阵对图像进行亚采样。
- 对于一维信号，采用前文提到的随机不等间距的亚采样即可。



三、压缩感知的重建方法

- 如前文所述，CS的重建也就是求解欠定方程组y=ΘS的方法。
- 这是一个零范数（l0）最小化问题，是一个NP完全问题（没有快速解法的问题），因此往往转换成一范数（l1）最小化的求解，或者用一些近似估计的算法。

四、总结

- 什么是压缩感知:
  - 如果一个信号**在某个变换域是稀疏的**，那么就可以用一个**与变换基不相关的**观测矩阵将变换所得高维信号投影到一个低维空间上，然后通过求**解一个优化问题**就可以从这些少量的投影中以高概率重构出原信号。

- 压缩感知可以用这样一句话来表述:
  - **直接采集出一个JPEG**
  - ——之前图像压缩的方法是全采样之后再压缩，抛弃稀疏变换域中的一些小系数；而CS直接减少了采样点，采集完后、经过重建的图像，就是一副在某变换域稀疏的压缩图像，比如JPEG。

Accepted from Open Call

Sparse Mobile Crowdsensing: Challenges and Opportunities

Leye Wang, Daqing Zhang, Yasha Wang, Chao Chen, Xiao Han, and Abdallah M'hamed

Abstract

Sensing cost and data quality are two primary concerns in mobile crowdsensing. In this article, we propose a new crowdsensing paradigm, *sparse mobile crowdsensing*, which leverages the spatial and temporal correlation among the data sensed in different sub-areas to significantly reduce the required number of sensing tasks allocated, thus lowering overall sensing cost (e.g., smartphone energy consumption and incentives) while ensuring data quality. Sparse mobile crowdsensing applications intelligently select only a small portion of the target area for sensing while inferring the data of the remaining unsensed area with high accuracy. We discuss the fundamental research challenges in sparse mobile crowdsensing, and design a general framework with potential solutions to the challenges. To verify the effectiveness of the proposed framework, a sparse mobile crowdsensing prototype for temperature and traffic monitoring is implemented and evaluated. With several future research directions identified in sparse mobile crowdsensing, we expect that more research interests will be stimulated in this novel crowdsensing paradigm.

urban data in regions that are not covered by the specialized sensing infrastructure.

To obtain high-quality sensed results in MCS applications, a straightforward idea is to recruit enough participants so as to ensure that their sensed data can cover almost the whole target area. Nevertheless, this strategy may incur high sensing cost, including overall smartphone energy and network bandwidth consumption, as well as incentives paid to the participants by the organizer.

Thus, *data quality* and *sensing cost* have certain intrinsic conflicts in MCS. A lot of existing studies have endeavored to address this issue by minimizing the redundant number of allocated tasks (or recruited participants) under the quality requirement of full or high coverage of all the sub-areas in a city [1, 2]. However, the number of allocated tasks in these studies should be roughly equivalent to the total number of sub-areas, which may still incur very high sensing cost. Then a question arises: *is it possible to further reduce the sensing cost by only sensing a small number of sub-areas while still guaranteeing a satisfactory level of data quality for the whole target area?* To answer this question, we investigate a novel MCS paradigm where only a small part of city sub-areas are sensed by par-

Sensing cost and data quality are two primary concerns in mobile crowdsensing. The authors propose a new crowdsensing paradigm, sparse mobile crowdsensing, which leverages the spatial and temporal correlation among the data sensed in different sub-areas to significantly reduce the required number of sensing tasks allocated, thus lowering overall sensing cost while ensuring data quality.

Introduction

- **Background**: With the prevalence of rich-sensor equipped smartphones in recent years, **mobile crowdsensing (MCS)** has become a promising paradigm to facilitate urban sensing applications, such as *environment monitoring, traffic congestion detection, hotspot identification, and public information sharing*.

To obtain **high-quality** sensed results in MCS applications, a straightforward idea is to recruit enough participants so as to ensure that their sensed data can cover almost the whole target area.

⟺

Nevertheless, this strategy may incur high **sensing cost**, including overall smartphone energy and network bandwidth consumption, as well as incentives paid to the participants by the organizer.

- **Question**: is it possible to further reduce the sensing cost by only sensing a small number of sub-areas while still guaranteeing a satisfactory level of data quality for the whole target area?

Overview of mobile crowdsensing process

- **A novel MCS paradigm —— Sparse MCS:** only a small part of city sub-areas are sensed by participants, while the data of the rest of the sub-areas are inferred based on the sensed data.
- **Theoretical feasibility:** high spatio-temporal correlations exist in most urban data (e.g., air quality and noise). Such correlations provide the basis for high-quality missing data inference. Specifically, recent research progress in missing data inference algorithms, such as compressive sensing, could facilitate Sparse MCS to achieve high inferred data quality.

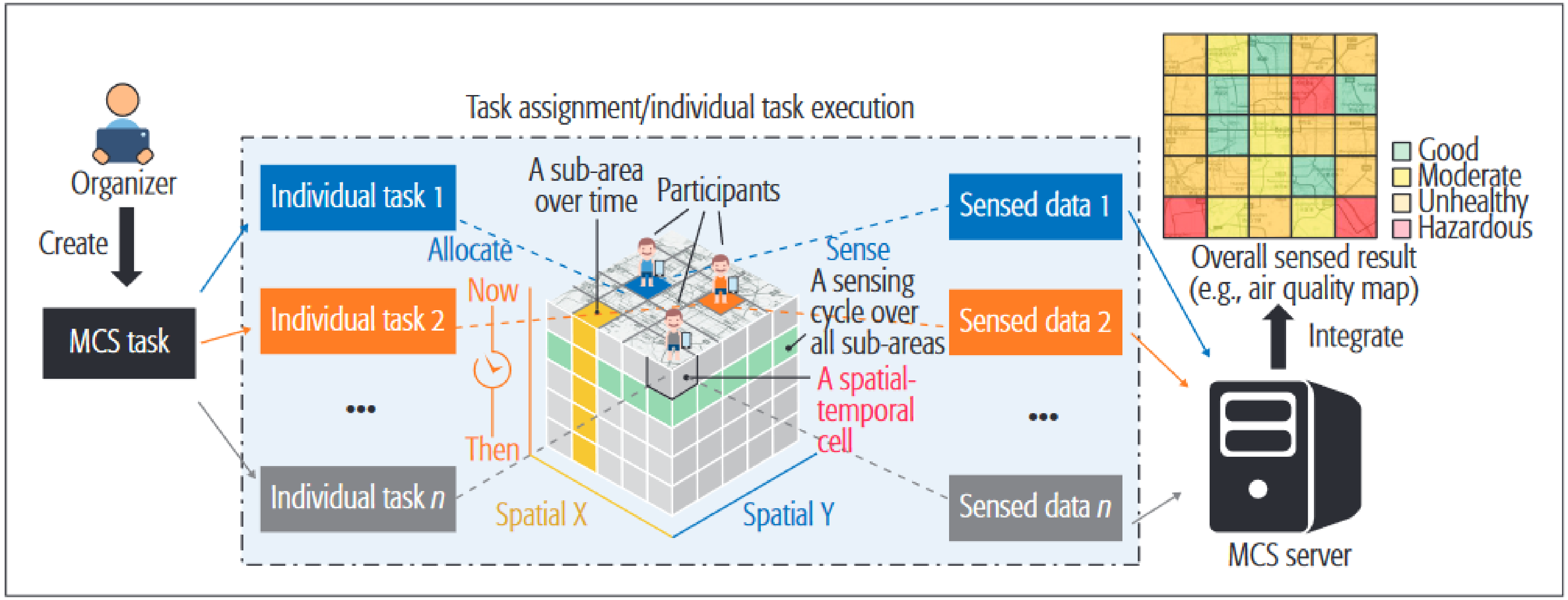


Figure 1. Overview of the mobile crowdsensing process. A cube is used to illustrate the complete set of all possible individual tasks, where each individual task is specified by a spatio-temporal cell (a specific sub-area in a specific cycle). Two dimensions (X and Y) of the cube represent the spatial space (sub-areas), and the other dimension (Z) represents the temporal space (cycles).

Key research challenges

- **Missing data inference**: Given a set of sensed data from the sparsely selected spatio-temporal cells, how do we infer the missing data of the remaining unsensed cells with high accuracy?
  - Different inference approaches have their own characteristics and applicable use cases; thus, the best inference algorithm may be **different** for heterogeneous Sparse MCS applications that have different data types, data sizes, and inherent correlations.
- **Optimal task allocation**: Given an inference algorithm, how do we select the optimal combination of spatio-temporal cells for task allocation so that the sensing cost is minimized with guaranteed inferred data quality?
  - different combinations of spatio-temporal cells will incur diverse inferred data quality and sensing cost.
  - the task allocation is a **monotonic and unregrettable** process —— if we have allocated tasks to some cells and collected the data, even if we find that the collected data is not efficient for improving the overall data quality, the allocation cannot be retracted to alter the previous decision.
  - we need to consider **user mobility** in task allocation to determine whether the selected cell can be covered by any participant or not
- **Data quality assessment**: Given a set of sensed spatio-temporal cells and an inference algorithm, how do we assess the inferred data quality without knowing the ground truth data values of unsensed cells?
  - the **lack of the ground truth** data of unsensed cells.

General framework for Sparse MCS

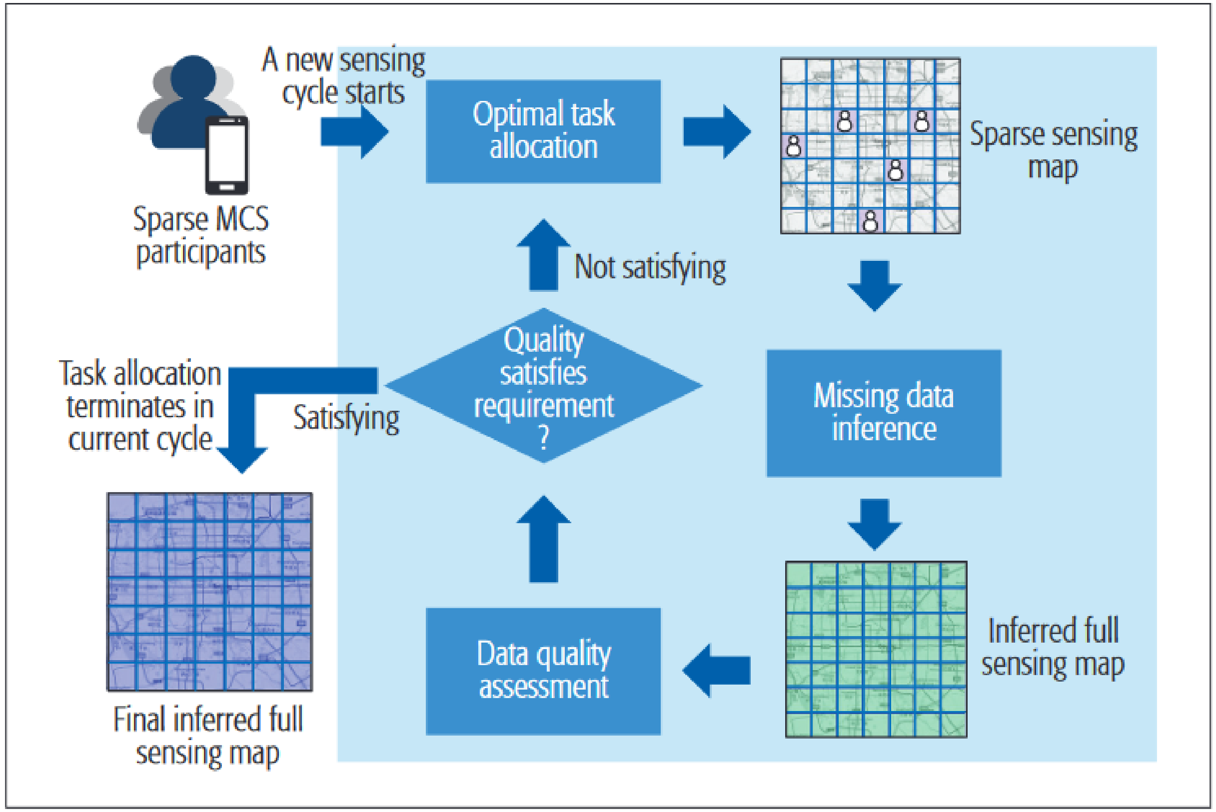


Figure 2. General framework for Sparse MCS applications in one sensing cycle.

- **Missing data inference**
  - compressive sensing, dynamic Bayesian network and Gaussian process regression
  - exploiting additional knowledge
  - multi-source data fusion algorithm
- **Optimal task allocation**
  - estimate the *inference uncertainty* of each unsensed cell
  - **Query by committee:** applies various algorithms to infer the value of an unsensed cell, and then calculates the variance among these inferred values as the uncertainty for that cell
  - the largest uncertainty is then selected as the next cell for sensing
- **Data quality assessment**
  - statistical techniques: re-sampling methods such as leave-one-out or bootstrap
  - the estimated inference error is the difference between the current-cycle inferred data and the last-cycle sensed data