Enhance the Quality of Crowdsensing for Fine-Grained Urban Environment Monitoring via Data Correlation

Xu Kang, Liang Liu * and Huadong Ma

Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia,
Beijing University of Posts and Telecommunications, Beijing 100876, China; kangxu88@bupt.edu.cn (X.K.);
mhd@bupt.edu.cn (H.M.)
* Correspondence: liangliu@bupt.edu.cn; Tel.: +86-10-6228-2767

Abstract: Monitoring the status of urban environments, which provides fundamental information for a city, yields crucial insights into various fields of urban research. Recently, with the popularity of smartphones and vehicles equipped with onboard sensors, a people-centric scheme, namely "crowdsensing", for city-scale environment monitoring is emerging. This paper proposes a data correlation based crowdsensing approach for fine-grained urban environment monitoring. To demonstrate urban status, we generate sensing images via crowdsensing network, and then enhance the quality of sensing images via data correlation. Specifically, to achieve a higher quality of sensing images, we not only utilize temporal correlation of mobile sensing nodes but also fuse the sensory data with correlated environment data by introducing a collective tensor decomposition approach. Finally, we conduct a series of numerical simulations and a real dataset based case study. The results validate that our approach outperforms the traditional spatial interpolation-based method.

Keywords: crowdsensing; urban sensing; environment monitoring

1. Introduction

Rapid urbanization makes the issue of urban environment more serious than ever before in major cities, especially for developing countries. Urban environment monitoring, which provides crucial information for scientific city management, is of great importance for solving urban environmental problems. *Urban Remote Sensing* (URS) [1] and specialized *Wireless Sensor Network* (WSN) [2,3] are commonly used ways for acquiring a wide range of environmental monitoring data. However, the principle of URS limits the precision for fine-grained urban sensing. While for WSN, the sensing nodes always face the limited energy [4] and instability [5] problem, which restricts the scalability and utility of WSN for urban sensing. Therefore, it is very necessary to develop innovative technologies that could precisely and ubiquitously sense urban dynamics.

Recent years, with the popularity of smartphones/vehicles equipped with onboard sensors and the development of wireless networks (such as 4G/Wi-Fi network), a people-centric sensing mode is emerging [6,7]. Researchers characterize this sensing mode as *crowdsensing* [8]. For urban monitoring with a crowdsensing network, mobile smartphones/vehicles equipped with sensing modules in a region send their measurements (sensory data) to a data center via wireless network, then the data center aggregates these measurements to estimate the phenomenon values of all points in the region. Similar to digital images, the distribution status of environment phenomenon (such as PM2.5 concentration, CO_2 concentration) over a region can be expressed as a two-dimensional (2D) signal. To illustrate fine-grained urban status, the crowdsensing network utilize sensory data to generate a

sensing image according to spatial interpolation methods [9]. As shown in Figure 1a,b, the crowdsensing network works like an "urban camera", which could capture the urban status.
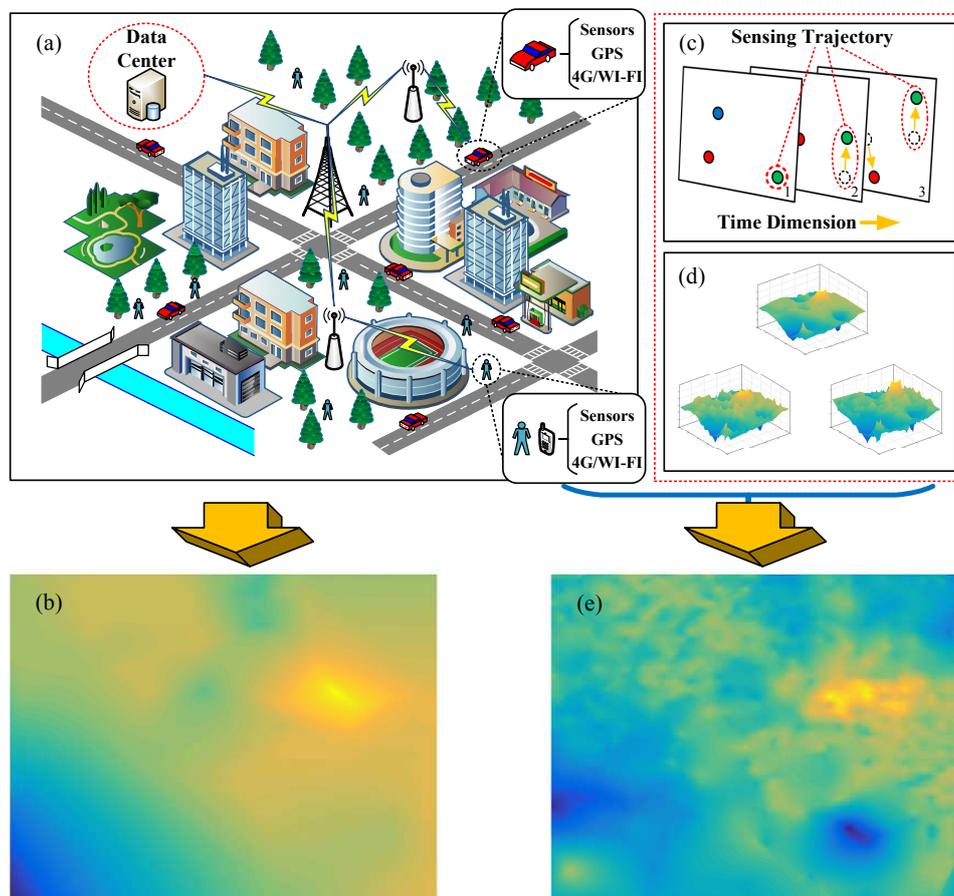
Figure 1. Crowdsensing networks based urban environment monitoring. (**a**) The crowdsensing networks, consisting of vast smartphones/vehicles, work as an urban camera; (**b**) Sensing image of target phenomenon generated by the crowdsensing network via interpolation method; (**c**) Sensing trajectories; (**d**) Correlated signals; (**e**) Sensing image after using enhanced crowdsensing approach.

Generally, the geographical positions of crowdsensing participants are dynamic and uneven distributed in an urban area, which makes the quality of sensing image varies by time and locations significantly. In our previous work [10], we first defined a metric – *urban resolution* – for measuring the quality of sensing image. Learning from the concept of resolution in conventional image system, higher/lower urban resolution of sensing image means more/less details the sensing image holds. In fact, this metric is a measurement for evaluating the sensing ability (sensitivity) of crowdsensing network. Furthermore, we also revealed a linear relationship between urban resolution r and the number of crowdsensing participants s from statistical perspective.

Intuitively, huge number of crowdsensing participants will provide a guarantee of high urban resolution. In practice, regardless of the number of crowdsensing participants, the spatial distribution of participants is uneven, i.e., there exist some hot zones and blank zones simultaneously in the urban area. This inhomogeneity is determined by the underlying human mobility features, which are summarized from real mobility traces [11]. Based on our observation, the emergence of blank regions will obviously reduce the sensing ability of the crowdsensing network. Thus, how to increase the sensing ability for blank zones becomes the main challenge.

Recently, many researchers devote to breaking the primary bottleneck of crowdsensing, i.e., sensory data sparsity, which is mainly caused by limited number or uneven distribution of

crowdsensing participants [12,13]. A powerful and generic technique, *Compressive Sensing* (CS), is extensively used for inferring the missing sensory data. By far, the application of CS has led to significant advances in reconstructing network traffic [14], improving urban traffic sensing [15] and refining localization [16], etc. Among these works, the CS-based methods share the same prerequisite that the sensory data should have inherent structure and can be sparsely represented. However, the CS-based method cannot achieve good performance for some common types of crowdsensing data (e.g., environment data), because the real environment data rarely satisfy the aforementioned technical conditions [17–19]. Fortunately, several works [17,20,21] have revealed both the temporal-spatial correlation and category correlation of environmental phenomenons through extensive real dataset analysis. These works motivate us to enhance the estimation accuracy of missing values for environment crowdsensing via a data correlation based approach. Here, the data correlation mainly refers to the following two aspects:

- Temporal correlation of sensory data. Researchers of [17] revealed the pervasive existence of time stability feature among environmental phenomenons, such as temperature, humidity and light. This feature indicates that most environmental phenomenons will not change dramatically and maintain stable for a while. On the other hand, the frequency of participants sending their measurements is much higher than that of environment changing. Utilizing temporal correlation is to leverage measurement data in a correlated time period rather than a moment. Due to the dynamic feature of crowdsensing participants, a discrete participant at a moment is converted into a sensing trajectory in correlated time period (see Figure 1c), which could decrease the area of blank zones.
- Category correlation of sensory data. Many existing researches [20–22] show the strong correlation among some categories of sensory data (see Figure 1d). Taking air quality data for example, three mainly concerned atmospheric pollutants, the concentration of PM2.5, PM10, and NO_2, have clear correlation. Therefore, if there exist some correlated sensory data in blank zones, then the correlated information is able to recover the target environmental phenomenon.

In this paper, we first express a sensing image as the form of matrix (discrete signal matrix) and utilize a Delaunay triangulation based interpolation to complete the matrix as a baseline. Then, we build a three-dimension tensor (each dimension denotes region, signal category and time slot respectively) to model measurements collected by participants, and summarize a signal correlation matrix to measure the correlation between different categories of environment phenomenons. After that, we propose a collaborative tensor decomposition approach to supplement the missing entries on the basis of temporal and category correlation of sensory data. In order to verify the enhancement of crowdsensing ability, we utilize the *Self-similar Least Action Walk* (SLAW) model [11], a commonly used human/vehicles mobility model to generate trajectories of crowdsensing participants. Then, we reconstruct a CO_2 concentration signal as the ground truth of target environment phenomenon, and further generate correlated signals through liner/non-liner functions. By conducting a series of simulations, we verify the promotion of crowdsensing ability. In addition, we also utilize real urban air pollutants data to examine our enhanced crowdsensing approach.

The rest of the paper is organized as follows. Section 2 presents the method for generating sensing image and defines crowdsensing resolution. Section 3 proposes the enhanced crowdsensing approach. Section 4 utilizes numerical simulations to investigate the resolution promotion of enhanced crowdsensing approach. Section 5 illustrates a case study by utilizing real dataset of urban air pollutants. The paper concludes in Section 6.

2. Urban Monitoring via Crowdsensing and Sensing Restriction of Crowdsensing

In this section, we first briefly introduce the process of generating sensing image via crowdsensing networks. Then, we give the definition of crowdsensing resolution, which is utilized to quantize

the sensing ability of crowdsensing networks. At last, we give the linear restriction of crowdsensing resolution, which limits the sensing ability of crowdsensing networks.

2.1. Generation of Sensing Image via Crowdsensing Network

For crowdsensing based environmental monitoring applications, the participants, i.e., mobile smartphones/vehicles equipped with sensing modules, are the sensing nodes. We define this kind of sensing nodes as *crowdsensing node*.

Definition 1. *Crowdsensing Node. A smartphone/vehicle with sensing and communication abilities is called a crowdsensing node, denoted by u_i, if the smartphone/vehicle senses the natural phenomenon. The value of u_i is represented as $u_i(x_i, y_i, v_i, t_i)$ where (x_i, y_i), v_i and t_i denote the position, measurement and the corresponding time of u_i, respectively.*

The natural phenomenon in a city is dynamic, while for a small time interval denoted by T, it can be expressed as a static two-dimensional (2D) signal. Mathematically, the signal is a function of two independent variables, i.e., spatial position (x, y) and the phenomenon value is $v = F(x, y)$. Figure 2a shows a case of 2D continuous-space signal (distribution of CO_2 concentration over a region) in T. Assume that the urban region is a square R and s crowdsensing nodes are dynamically distributed in R. Figure 2b illustrates the distribution of 100 crowdsensing nodes at time $t_i \in T$. Let $U = \{u_i, 1 \leq i \leq s\}$ be the set of crowdsensing nodes in R, and $V = \{v_i, 1 \leq i \leq s\}$ be the set of measurements. In other words, V is the set of sampling points of $F(x, y)$ (see Figure 2c). Generation of sensing image, is to estimate the phenomenon values of all points in R via V (see Figure 2d).

(a)   (b)

(c)   (d)
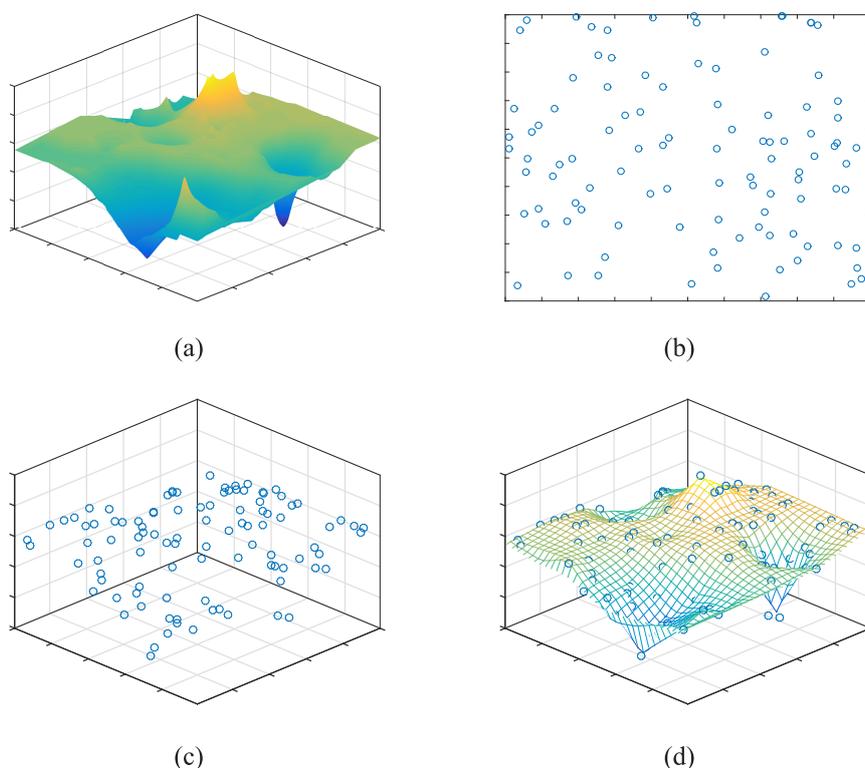
Figure 2. The process of generating sensing image via crowdsensing. (**a**) The 2D signal of CO_2 concentration; (**b**) The distribution of 100 nodes in R with their sending time $t_i \in T$; (**c**) The sampling points of 2D signal; (**d**) Sensing image generated by V.

In order to implement this estimation, we transform $F(x, y)$ from 2D continuous-space signal to a 2D discrete-space signal by dividing the unite square R into $m \times m$ grids (see Figure 3a). Each grid has a uniform signal value. Then, the 2D signal could be expressed by a matrix $\mathbf{Z}_{m \times m}$ for time period T. Because of the uneven distribution of crowdsensing nodes, which is determined by human mobility, there exists three situations for a given grid $g_{i,j}$:

- There is only one crowdsensing node in $g_{i,j}$. For this case, the corresponding entry $z_{i,j}$ in matrix $\mathbf{Z}$ is equal to the sensory data provided by the only crowdsensing node.
- There are more than one crowdsensing nodes in $g_{i,j}$. The corresponding $z_{i,j}$ is calculated by a weight sum of all the sensory data generated in $g_{i,j}$. Considering the distribution of crowdsensing nodes in one grid, we build a voronoi diagram according to the locations of crowdsensing nodes (see Figure 3b), and then calculate the weight sum of sensory data where the area of the divided polygons are weights of sensory data.
- There is no crowdsensing node in $g_{i,j}$. The corresponding $z_{i,j}$ is set to null.

(a)   (b)
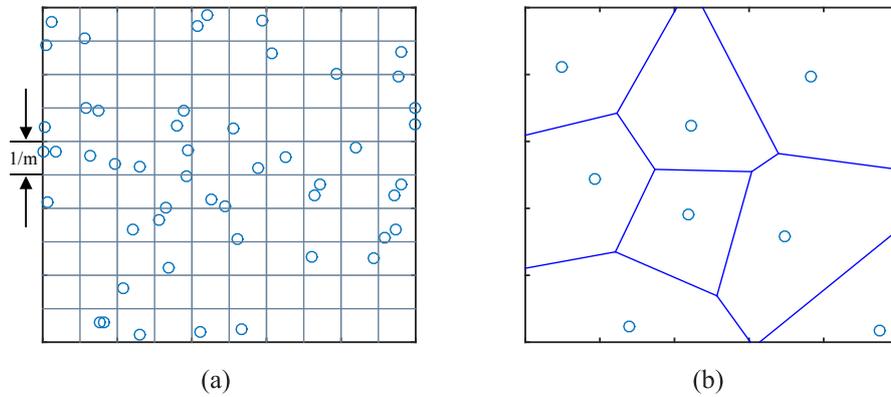
Figure 3. (**a**) Division of unit square R by $m \times m$ uniform grids; (**b**) A voronoi diagram based on the locations of crowdsensing nodes.

A commonly used way for filling the null entries in $\mathbf{Z}$ is spatial interpolation method. More details about spatial interpolation can be found in [9]. Based on the interpolation method, we give the definition of sensing image as follows:

Definition 2. *Sensing Image. For a given time interval T, the sensing image generated by the set of crowdsensing nodes, U, is defined as the following 2D matrix:*

$$\mathbf{Z}' = \begin{bmatrix} z'_{1,1(T)} & z'_{1,2(T)} & \cdots & z'_{1,m(T)} \\ z'_{2,1(T)} & z'_{2,2(T)} & \cdots & z'_{2,m(T)} \\ \vdots & \vdots & \ddots & \vdots \\ z'_{m,1(T)} & z'_{m,2(T)} & \cdots & z'_{m,m(T)} \end{bmatrix}_{m \times m,}$$

where

$$z'_{i,j(T)} = \begin{cases} \sum_{u_k \in U_{i,j}} v_k \frac{a_k}{g}, & |U_{i,j}| > 0; \\ I_n\left(V', g_{i,j}\right), & |U_{i,j}| = 0. \end{cases} \tag{1}$$

In Equation (1), $U_{i,j}$ denotes the set of crowdsensing nodes in grid $g_{i,j}$ within T. Besides, g, a_k are the area of the whole grid $g_{i,j}$ and the polygon centered on u_k in the corresponding voronoi diagram respectively. $I_n(\cdot, \cdot)$ denotes the interpolation method and $V' \triangleq \left\{ z'_{i,j(T)} : \forall g_{i,j}, |U_{i,j}| > 0 \right\}$.

2.2. Resolution of Crowdsensing

As mentioned in Section 1, to measure the ability of crowdsensing is to estimate the quality of sensing image, i.e., *QoI*. Firstly, we give the definition of *QoI* as follows:

Definition 3. *Quality of sensing Image. The quality of a sensing image for a specific time interval T, is defined as the similarity between the sensing image* $\mathbf{Z}'$ *and the static 2D discrete-space raw signal* $\mathbf{Z}$ *of the target environment phenomenon. Correlation coefficient is used for measuring the similarity, i.e.,*

$$C(\mathbf{Z}', \mathbf{Z}) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{m}\left(z_{i,j(T)} - \overline{z}_{i,j(T)}\right)\left(z'_{i,j(T)} - \overline{z}'_{i,j(T)}\right)}{\sqrt{\left(\sum_{i=1}^{m}\sum_{j=1}^{m}\left(z_{i,j(T)} - \overline{z}_{i,j(T)}\right)^2\right)\left(\sum_{i=1}^{m}\sum_{j=1}^{m}\left(z'_{i,j(T)} - \overline{z}'_{i,j(T)}\right)^2\right)}}.$$

A large correlation value means a strong similarity between $\mathbf{Z}'$ and $\mathbf{Z}$, i.e., the QoI of $\mathbf{Z}'$ is high. However, for most cases, the raw signal $\mathbf{Z}$ is unknown in advance. Therefore, users cannot calculate the QoI by using this equation directly. Then, we turn to use a new metric, *sensing image resolution* (i.e., urban resolution proposed in [10]), to measure the quality of sensing image $\mathbf{Z}'$.

In conventional image systems, *resolution*, defined as the number of gridded pixels, is a metric to evaluate the quality of image. Similarly, resolution can also be used to measure the quality of sensing image. However, the challenge is that the resolution of sensing image is not simply the crowdsensing nodes' count. As the crowdsensing nodes are dispersedly and dynamically distributed and not the strictly gridded distribution.

Indirectly, if we could find out how many gridded sensing nodes are needed to generate a sensing image, denoted by $\mathbf{Z}''$, which has the similar QoI with $\mathbf{Z}'$ generated by s crowdsensing nodes, then the number of gridded sensing nodes, $n \times n$, can be regarded as the resolution of sensing image $\mathbf{Z}'$. On the other hand, the corresponding $n \times n$ is also an estimation for the sensing ability of crowdsensing network. Therefore, we redefine this metric as *crowdsensing resolution*.

Definition 4. *Resolution of Crowdsensing. The resolution, denoted by r, of a crowdsensing network with s sensing nodes is defined as* $n_l \times n_l$ *where*

$$n_l = \arg\max_{n} C\left(\mathbf{Z}', \mathbf{Z}''\right). \tag{2}$$

In Equation (2), $\mathbf{Z}'$ denotes the sensing image generated by the crowdsensing network, and $\mathbf{Z}''$ denotes a sensing image generated by $n \times n$ gridded sensing nodes.

Obviously, when n_l becomes large enough, $\mathbf{Z}''$ could approximately equal to $\mathbf{Z}$, i.e., $C\left(\mathbf{Z}'', \mathbf{Z}\right) \approx 1$. That is to say, the crowdsensing network is with high sensing ability.

2.3. Linear Restriction

In paper [10], we utilized three 2D signals with different variation degrees to analyze the relationship between crowdsensing resolution r and sensing nodes number s via Monte Carlo simulations. An approximately linear growth relationship between r and s is revealed.

For s crowdsensing nodes in an unit area, the relationship between n_l and s follows:

$$n_l = \alpha\sqrt{s},$$

where the slope α is with the reference value in $[0.5, 0.6]$.

Therefore, given the number of crowdsensing participants s, we can easily infer the resolution r of the crowdsensing network via the linear relationship. On the other hand, the linear relationship

performs as a restriction, which limits the sensing ability that a crowdsensing network could achieve according to the scale of crowdsensing participants.

3. Enhanced Crowdsensing Approach

To enhance the crowdsensing ability and break the linear restriction, we utilize both signal and temporal correlation of sensory data to generate sensing images more precisely. Firstly, we introduce how to model the sensing data acquired by crowdsensing participants. Then, we propose a collaborative tensor decomposition approach to infer the missing items of target signal through signals' correlation. At last, we illustrate the correlated time slots combination approach to further infer the rest missing items via temporal correlation of environmental signal.

3.1. Data Modeling

As shown in the left part of Figure 4, we model the sensory and correlated sensory data using a tensor, $\mathbf{S} \in \mathbb{R}^{N \times Q \times K}$, with three dimensions:

- *Region dimension*, $\mathbf{r} = [r_1, r_2, \cdots, r_N]$, denotes N regions which are transferred from $m \times m$ grids, one region per grid, $g_{i,j} = r_{im+j}$.
- *Category dimension*, $\mathbf{c} = [c_1, c_2, \cdots, c_Q]$, denotes Q signal categorise, where c_1 is the target signal and the others are correlated signals.
- *Time dimension*, $\mathbf{t} = [t_1, t_2, \cdots, t_K]$, denotes K time slots. Here, we divide the monitoring time period into K time slots and the span of each time slot is decided by the sending intervals of crowdsensing nodes.

An entry $\mathbf{S}_{i,j,k}$ stores the sensory data, which is acquired in region r_i at time slot t_k with signal category of c_j. Likewise, for the process of mapping data to the tensor $\mathbf{S}$, there also exists three situations as mentioned in Section 2.1 for a specific region and we adopt the same strategy to calculate the sensory data.
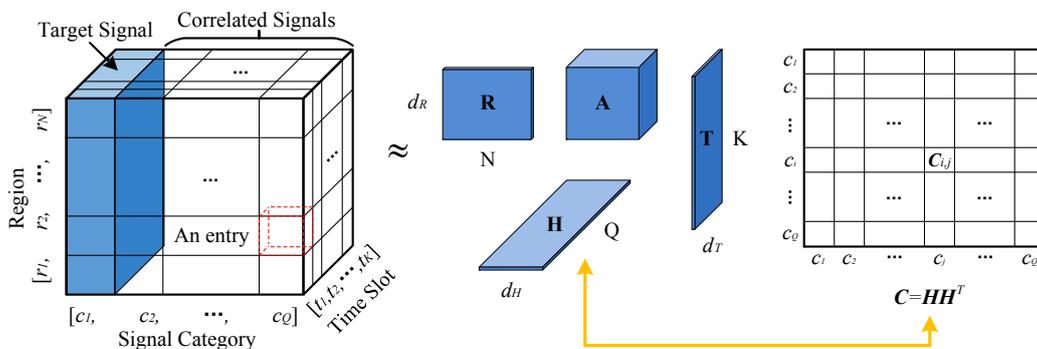
Figure 4. Collaborative tensor decomposition.

As tensor $\mathbf{S}$ is sparse, we try to "borrow" more information from correlation between different categories of signals for inferring the missing entries. Although tensor $\mathbf{S}$ can capture the correlation between different categories of signals to some context, a specialized matrix can further intensify the correlation. As shown in the right part of Figure 4, we formulate a matrix $\mathbf{C} \in \mathbb{R}^{Q \times Q}$ to model the signals' correlation, where an entry $\mathbf{C}_{i,j}$ denotes the correlation between signal c_i and c_j. The correlation is quantized by:

$$\mathbf{C}_{i,j} = \frac{\sum\limits_{t}\sum\limits_{r} \left(\mathbf{S}_{r,i,t} - \bar{\mathbf{S}}_{r,i,t}\right)\left(\mathbf{S}_{r,j,t} - \bar{\mathbf{S}}_{r,j,t}\right)}{\sqrt{\left(\sum\limits_{t}\sum\limits_{r}\left(\mathbf{S}_{r,i,t} - \bar{\mathbf{S}}_{r,i,t}\right)^2\right)\left(\sum\limits_{t}\sum\limits_{r}\left(\mathbf{S}_{r,j,t} - \bar{\mathbf{S}}_{r,j,t}\right)^2\right)}}, \tag{3}$$

where $\sum\limits_t$ is $\sum_{t=1}^{K}$ and $\sum\limits_r$ is $\sum_{r=1,\mathbf{S}_{r,i,t}\neq null,\mathbf{S}_{r,j,t}\neq null}^{N}$.

After data modeling, we get a sparse tensor $\mathbf{S}_{N\times Q\times K}$ and a signal correlation matrix, $\mathbf{C}_{Q\times Q}$.

3.2. Collaborative Tensor Decomposition

To achieve a higher accuracy of filling the missing entries in tensor $\mathbf{S}$, we exploit the collaborative tensor decomposition method. As shown in the middle part of Figure 4, tensor $\mathbf{S}$ is transformed into tucker decomposition model [23], i.e., the multiplication of a core tensor $\mathbf{A} \in \mathbb{R}^{d_R\times d_H\times d_T}$ with three matrices, $\mathbf{R} \in \mathbb{R}^{N\times d_R}, \mathbf{H} \in \mathbb{R}^{Q\times d_H}, \mathbf{T} \in \mathbb{R}^{K\times d_T}$. Here, d_R, d_H, d_T are very small, denoting the number of latent factors. Moreover, the signal correlation matrix $\mathbf{C}$ is decomposed as the self product of matrix $\mathbf{H}_{Q\times d_H}$, where $d_H < Q$, by low-rank approximation. By requiring them share the same low rank matrix $\mathbf{H}_{Q\times d_H}$, we could propagate the information between tensor $\mathbf{S}$ and matrix $\mathbf{C}$. Then, we define a loss function to quantize the error of collaborative tensor decomposition as:

$$
\begin{aligned}
\Gamma\left(\mathbf{A},\mathbf{R},\mathbf{H},\mathbf{T}\right) = &\frac{1}{2}\left\|\mathbf{I}\circ\left(\mathbf{S}-\mathbf{A}\times_R\mathbf{R}\times_H\mathbf{H}\times_T\mathbf{T}\right)\right\|^2 + \frac{\lambda_1}{2}\left\|\mathbf{C}-\mathbf{H}\mathbf{H}^T\right\|^2 \\
&+ \frac{\lambda_2}{2}\left(\|\mathbf{A}\|^2+\|\mathbf{R}\|^2+\|\mathbf{H}\|^2+\|\mathbf{T}\|^2\right),
\end{aligned}
\tag{4}
$$

where $\|\cdot\|$ denotes the Frobenius norm, $\mathbf{I}$ is an indicating tensor, and the operator "$\circ$" denotes the entry-wise product. The entry $\mathbf{I}_{i,j,k} = 0$ if $\mathbf{S}_{i,j,k}$ is missing, $\mathbf{I}_{i,j,k} = 1$ otherwise. In Equation (4), $\left\|\mathbf{I}\circ\left(\mathbf{S}-\mathbf{A}\times_R\mathbf{R}\times_H\mathbf{H}\times_T\mathbf{T}\right)\right\|^2$ and $\left\|\mathbf{C}-\mathbf{H}\mathbf{H}^T\right\|^2$ are to measure the errors of decomposing $\mathbf{S}$ and $\mathbf{C}$, respectively, and the last part $\|\mathbf{A}\|^2 + \|\mathbf{R}\|^2 + \|\mathbf{H}\|^2 + \|\mathbf{T}\|^2$ is regularization penalty to avoid over-fitting. Parameters λ_1 and λ_2 are to control the contribution of each part for collaborative decomposition.

Minimize the loss function above, is to find an optimal result for inferring the missing entries in tensor $\mathbf{S}$ with data correlation. But this loss function is not jointly convex to all the variables of $\mathbf{A},\mathbf{R},\mathbf{H}$ and $\mathbf{T}$. In general, we cannot get closed-form solutions for the minimization. Therefore, we utilize a gradient descent technique [21] to get a local optimal solutions by minimizing the loss function iteratively. Finally, we recover $\mathbf{S}$ by $\mathbf{S}_{rec} = \mathbf{A}\times_R\mathbf{R}\times_H\mathbf{H}\times_T\mathbf{T}$. Here, the multiplication symbol with subscript denotes tensor mode multiplication, e.g., $\mathbf{X} = \mathbf{A}\times_R\mathbf{R}$ is $\mathbf{X}_{n,j,k} = \sum_{i=1}^{d_R}\mathbf{A}_{i,j,k}\times\mathbf{R}_{n,i}$.

The mobility features of human crowd result in the uneven distribution of crowdsensing nodes, therefore it forms several hot/blank regions in urban area. For hot regions, there exist many crowdsensing nodes providing their measurements. As a result, different categories of sensory data co-exist in these regions, it is helpful for mining the signal correlation. On the other hand, there also exist some blank regions without any sensory data both for target signal and correlated signals. This means in such regions the target signal can't be precisely recovered simply by collaborative tensor decomposition.

3.3. Correlated Time Slots Combination

As mentioned in Section 1, the frequency of crowdsensing nodes sending their measurements is much higher than that of environment changing, and the environment phenomenon will keep steady in a relative long period. To further supplement the missing items, we combine several continuous time slots together to generate a sensing image. For a specific crowdsensing node, correlated time slots combination is to make extension from time dimension. Due to the dynamic feature of crowdsensing nodes, in a period, these scattered crowdsensing nodes are more likely to form several sensing traces, therefore cover more regions.

Given a time period $T \in [t_s, t_e]$, we extract the corresponding slices in tensor $\mathbf{S}$ and make an entry-wise combination (see Figure 5). Here, a specific extracted slice for time slot t_i is a region-signal

matrix. After that we formulate a matrix **M** (see the right part of Figure 5). An entry $m_{i,j}$ in matrix **M**, is computed by:

$$m_{i,j} = \begin{cases} \dfrac{\sum\limits_{t_s \leq t \leq t_e} \mathbf{S}(i,j,t)}{N}, & N \neq 0; \\ null, & N = 0, \end{cases}$$

where N denotes the number of sensory data with $\mathbf{S}(i,j,t) \neq null$ in $[t_s, t_e]$. The first column of **M**, denoted by $V_1 = [m_{1,1}, m_{2,1}, \cdots, m_{N,1}]^T$, is the sensory data of target signal in N regions. By converting V_1 back into $m \times m$ matrix and using spacial interpolation method to further recover the missing entries, we get the optimized sensing image.
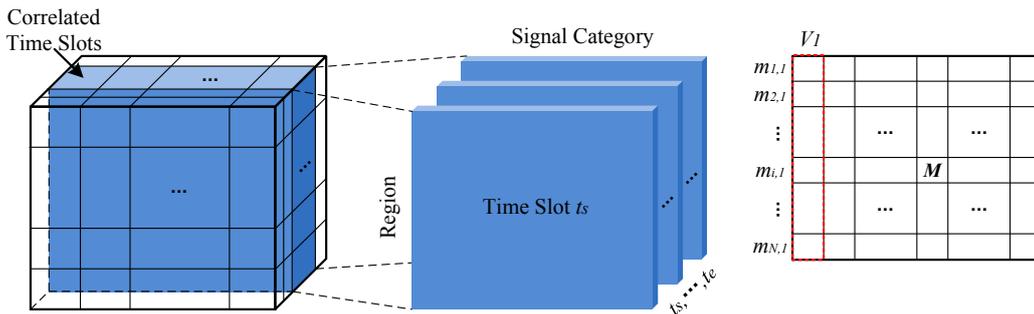
Figure 5. Correlated time slots combination.

4. Numerical Simulation

In this section, we conduct numerical simulations to examine the promotion of our crowdsensing enhancement approach. Firstly, we illustrate the models we used to generate the mobile traces and environment signals. Then, we demonstrate an instance of the generated sensing images via enhanced crowdsensing approach. At last, we verify the promotion of our approach statistically.

4.1. Basic Models

(1) Mobility model of crowdsensing nodes: Existing works about human/vehicle mobility have summarized the following statistical features from real mobile traces: **F1** *heterogeneous bounded mobility areas*, **F2** *truncated power-law flights and pause-times*, **F3** *truncated power-law inter-contact times*, and **F4** *fractal waypoints*. In this paper, we utilize *Self-similar Least Action Walk* (SLAW) [11], that produces synthetic walk trajectories containing all above features, to simulate mobility traces of crowdsensing nodes. Parameter settings are summarized in Table 1. More details about SLAW model and the interpretation of these parameters can be found in [11].

Table 1. Parameter settings of SLAW model.

Parameter	Value
Distance alpha	3
Number of mobile nodes	6000
Simulation area	$5000 \times 5000 \, \mathrm{m}^2$
Number of waypoints	6000
Hurst parameter	0.75
Time duration	10 h
Clustering range	100 m
Levy exponent for pause time	1
Minimum/maximum pause time	30 s/1800 s

After executing the SLAW model, we obtain all the positions of 6000 crowdsensing nodes for every minute over 10 h. Figure 6, illustrates the distribution of the crowdsensing nodes over the whole simulation region at a given instant. From Figure 6, we observe that some hot zones and blank zones exist simultaneously. This implies that the sensing ability of crowdsensing network for different urban regions may be very different. On the other hand, it is extremely complex to generate a sensing image of the whole city via all crowdsensing nodes in the city. Therefore, we divide the simulation region into 25 unit grids with $1 \times 1\,\text{km}^2$, and study the crowdsensing resolution for each grid.
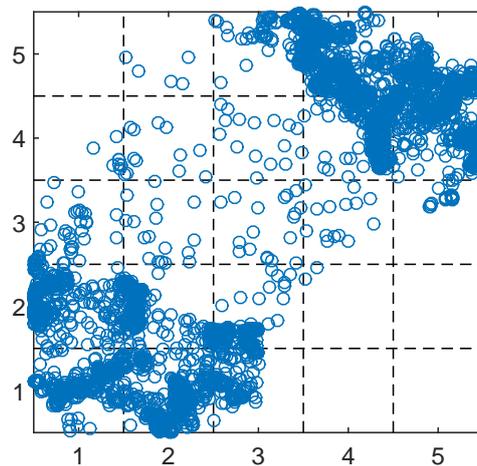
Figure 6. An instant of 6000 crowdsensing nodes generated by SLAW model.

(2) Target and correlated signal models: We utilize the 2D signal shown in Figure 2a as the ground truth of target signal. This is a signal of CO_2 concentration which is generated by the sensory data from 100 CO_2 sensor nodes in a square region (about $1\,\text{km}^2$) of Wuxi City, China [24]. To generate the correlated signals, we utilized linear/non-linear transform function as follows:

$$f_1\left(z_{i,j}\right) = az_{i,j} + b + e,$$
$$f_2\left(z_{i,j}\right) = az_{i,j}^2 + bz_{i,j} + c + e,$$

where $e \sim N(0, \delta)$ is the additive noise. By adjusting the parameters of a, b, c and δ, we generate the linear/non-linear signals with 0.9 correlation (as shown in Figure 7).

(a) (b)
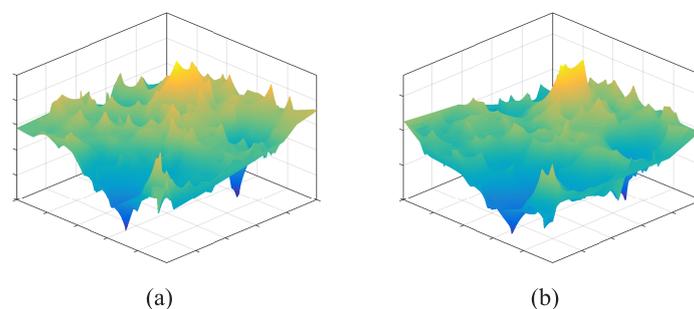
Figure 7. Correlated signals. (**a**) Linear correlated signal: $f_1\left(z_{i,j}\right) = 0.95 \times z_{i,j} + 0.05 + N\left(0, 0.06\right)$; (**b**) Non-linear correlated signal: $f_2\left(z_{i,j}\right) = 0.9 \times z_{i,j}^2 + 0.9 \times z_{i,j} + 0.1 + N\left(0, 0.11\right)$.

4.2. Instance Illustration

We randomly divide the 6000 crowdsensing nodes into 3 groups (2000 nodes for each group) (see Figure 8a). Group I senses the target signal, Group II and III sense the correlated signals, respectively. Assuming that, the correlated time period of environment phenomenon is one hour. To estimate how the frequency of crowdsensing nodes sending their measurements affects crowdsensing promotion, we conduct two experiments by adjusting the sending interval as 10/20 min. Here, we extract the positions of sensing nodes from SLAW traces every 10/20 min to simulate different sending intervals. Therefore, for the time duration of one hour, we have 6/3 correlated time slots respectively. We name these two experiments by S6 and S3 respectively.

(a)   (b)
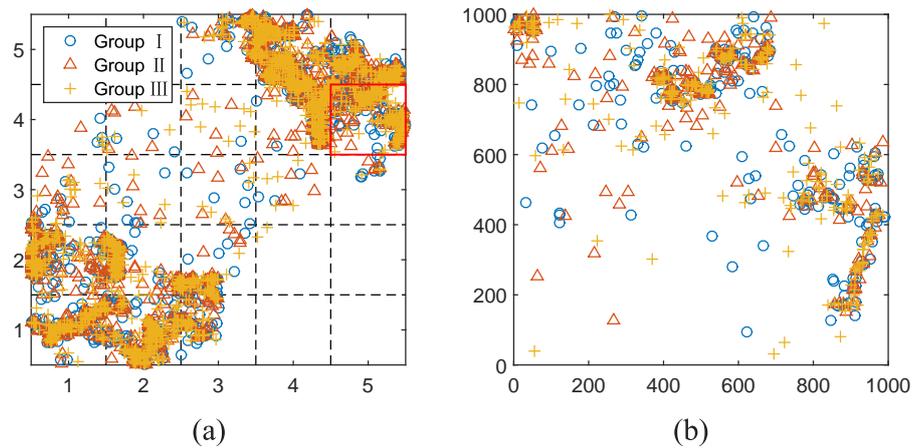
Figure 8. (**a**) Distribution of 6000 nodes (Groups I,II, and III) over the whole 5×5 km^2 region which is divided into 25 unit squares; (**b**) Distribution of crowdsensing nodes over the unit area indicated by the red box in (**a**).

As shown in Figure 9, for instance, we choose an unit grid (the red box in Figure 8a) to illustrate the recovered target signal after using enhanced crowdsensing approach. To make comparison, we also select a time slot in this one hour time period and generate a sensing image on the basis of interpolation method (as shown in Figure 9a).

(a)   (b)   (c)
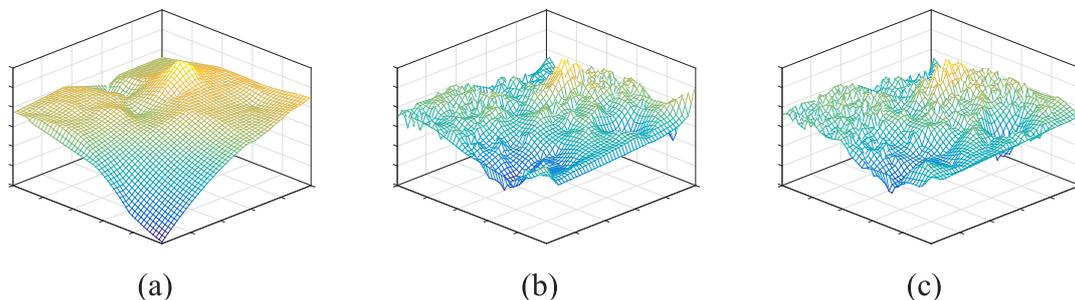
Figure 9. Recovered signals (generated sensing images). (**a**) Interpolation method; (**b**) Enhanced crowdsensing approach with 3 time slots combination; (**c**) Enhanced crowdsensing approach with 6 time slots combination.

The qualities of Figure 9a–c are 0.653, 0.884 and 0.929, respectively. In this unit grid, the average number of crowdsensing nodes (Group I) is 247. Then, we use $n \times n$ gridded sampling points to generate sensing images and compute $C\left(\mathbf{Z}'', \mathbf{Z}\right)$ against different values of n

(see Figure 10). Let $C(\mathbf{Z}', \mathbf{Z}) \approx C(\mathbf{Z}'', \mathbf{Z})$, we can easily acquire the corresponding n_l as 7.72, 12.57 and 13.48 respectively.
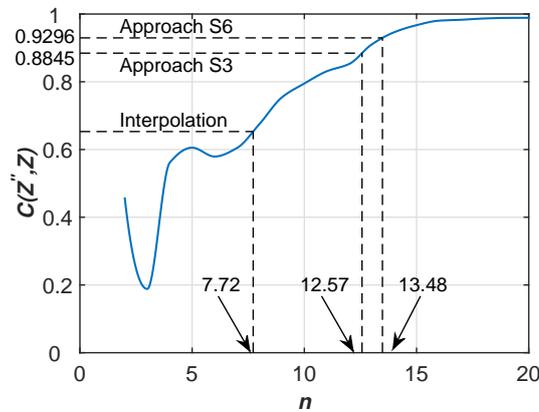
Figure 10. Correlation coefficients $C(\mathbf{Z}'', \mathbf{Z})$ against different values of n.

4.3. Statistical Results

In this section, we analyze the crowdsensing promotion from statistical perspective. We count the number of Group I nodes for each grid at every time slots and calculate the average number of nodes for every hour. For each grid, we recover the target signal for every hour based on crowdsensing promotion approach with different sending intervals (10/20 min). Then we calculate the corresponding n_l for each sensing image according to Equation (2). Therefore we have 25 grids $\times$ 10 h $\times$ 1 n_l values. We calculate the mean of n_l which are with the same $\lfloor \sqrt{s} \rfloor$. Figure 11 plots the n_l results against different values of $\sqrt{s}$. Moreover, we also plot the n_l result based on interpolation method.

According to the linear regression results of simulations, we obtain the linear function $n_l = 0.583\sqrt{s}$ for interpolation method. The slope α value of 0.583 is consistent with the result mentioned in paper [10]. We also get the linear functions: $n_l = 0.774\sqrt{s}$ and $n_l = 0.831\sqrt{s}$, for enhanced crowdsensing with S3 and S6 respectively. The higher value of α is, the higher resolution achieves for a given s. This statistical result validate that our enhanced crowdsensing approach outperforms the interpolation based method.
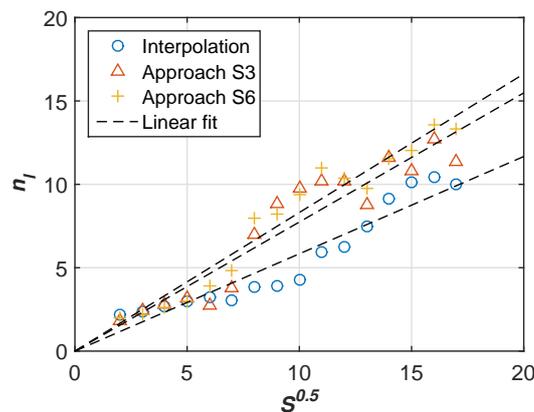
Figure 11. Relationship between crowdsensing resolution r and s crowdsensing nodes in an unit area which are generated by SLAW model. The x-axis denotes $\sqrt{s}$, and the y-axis denotes n_l, i.e., $\sqrt{r}$.

5. Case Study: Air Quality Monitoring in Beijing

The U-Air project [20] which is led by Microsoft Research aims to inference fine-grained air quality index (AQI) in Beijing with ubiquitous data. The researchers of U-Air open two data sources they used: locations of air monitoring stations in Beijing, and AQI (including the composition of air pollutants) of each station. In this section, we utilize a portion of these data to verify the effectiveness of our crowdsensing enhancement approach.

5.1. Data Correlation Analysis

In [20], researchers reveal that the concentration of air pollutants is notably influenced by meteorology (e.g., temperature, humidity) through analysing the reported information of monitoring stations. Figure 12a shows the locations of 22 monitoring stations in Beijing, and these stations report AQI every one hour. We extract the parameters of PM2.5, PM10 and air humidity index of monitoring station $S9$ within the time period from 21 to 27 April 2013, and quantize their correlation by correlation coefficients, denoted by $C(\cdot, \cdot)$.

(a)   (b)
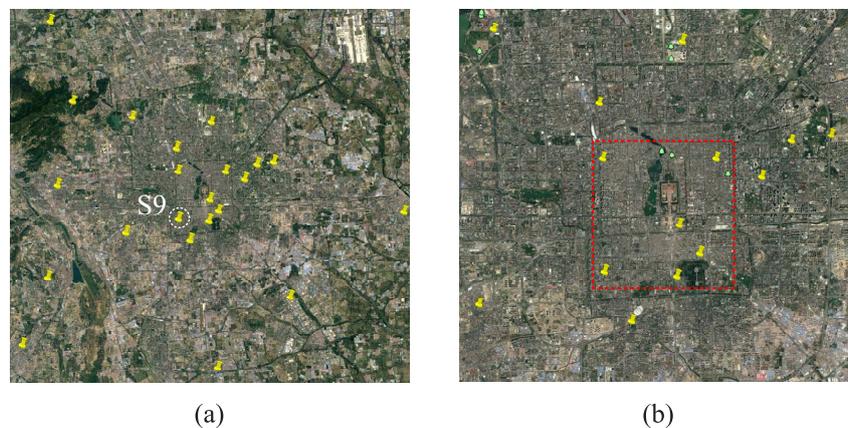
Figure 12. (**a**) Air quality monitoring stations in Beijing viewed from Google Earth; (**b**) The selected monitoring area within 2th ring road of Beijing.

Figures 13 and 14 shows the variation tendency of PM2.5 versus PM10 and air humidity index, respectively. Both of the figures reflect clear correlation among PM2.5, PM10 and humidity. Specifically, the $C(PM2.5, PM10)$, $C(PM2.5, Humidity)$ are 0.91 and 0.75. Moreover, we also analyse the correlation among PM2.5, PM10 and air humidity index for the rest 21 monitoring stations. The statistical results are given in Table 2.
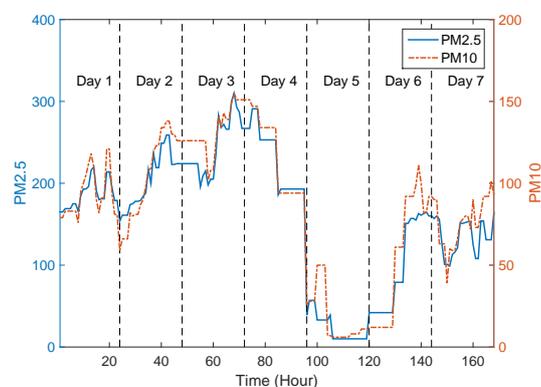
Figure 13. Variation tendency of PM2.5 and PM10 in monitoring station $S9$ during the time period of 21–27 April 2013.
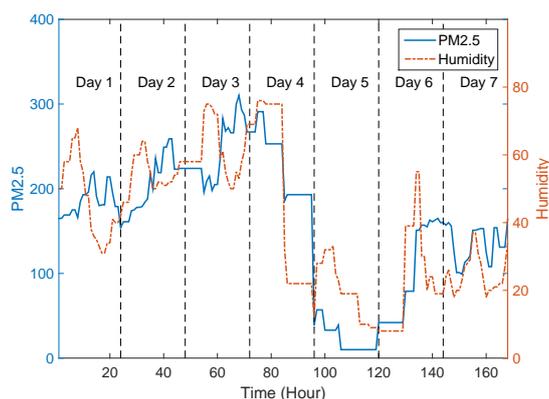
Figure 14. Variation tendency of PM2.5 and air humidity in monitoring station *S*9 during the time period of 21–27 April 2013.

From Table 2, we observe that there are 20 monitoring stations with their $C(PM2.5, PM10)$ above than 0.8, and 20 monitoring stations with their $C(PM2.5, Humidity)$ above than 0.7, among all the 22 monitoring stations. This statistical result reveals strong correlation among PM2.5, PM10 and air humidity index.

Table 2. Statistical result of correlation analysis.

Value Ranges of $C(\cdot, \cdot)$	Number of Monitoring Stations	
	$C(PM2.5, PM10)$	$C(PM2.5, Humidity)$
$C(\cdot, \cdot) \in [0.9, 1.0)$	13	0
$C(\cdot, \cdot) \in [0.8, 0.9)$	7	1
$C(\cdot, \cdot) \in [0.7, 0.8)$	1	19
$C(\cdot, \cdot) \in [0.5, 0.7)$	1	2

5.2. Data Modelling

Due to the lack of monitoring stations (Beijing only has 22 stations covering a 50×50 km^2 land, i.e., 113 km^2/per station [20]), it is not possible to acquire a fine-grained air pollutants distribution image for the whole urban area simply by spatial interpolation method. We choose the area within the second ring road of Beijing (see the red box zone in Figure 12b), which is the most concentrated area of monitoring station, as the target monitoring area. We map 3 categories of AQI (PM2.5, PM10 and air humidity index) onto the selected area according to the locations of their corresponding monitoring stations, and recover the signals by spatial interpolation method for time period around 12:00 at 24 April 2013 (see Figure 15).
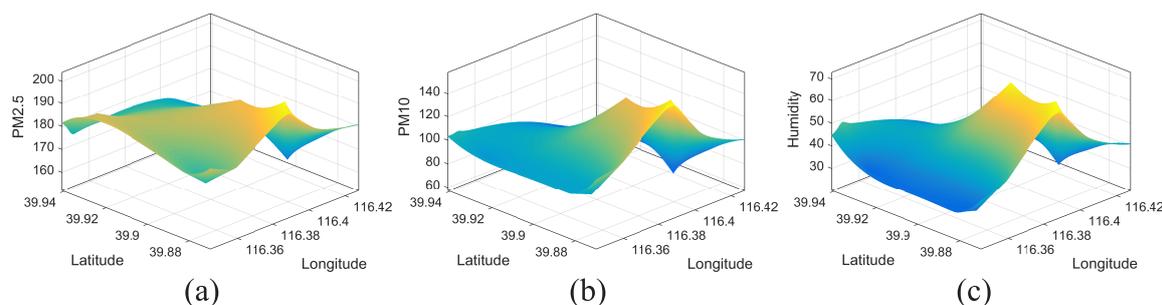
Figure 15. The constructed signals within the red box zones of Figure 12b. (**a**) Signal of PM2.5; (**b**) Signal of PM10; (**c**) Signal of air humidity.

From Figure 15, we observe similar distribution status among PM2.5, PM10 and air humidity. We take signal of PM2.5 as the target signal, while the others are correlated signals.

5.3. Comparison of Generated Sensing Images

Microsoft Research shared a collection of taxicab GPS traces gathered in Beijing [25], and these traces are perceived with the statistical features of human mobility. In our crowdsensing network, we deem these ordinary taxicabs as the crowdsensing nodes and randomly select 100 different taxi traces within the target zones. Then, we extract an instant of the taxis' GPS locations and sense the PM2.5 signal in Figure 15a. The generated sensing image by spatial interpolation method is shown in Figure 16a. Similar as mentioned in Section 4.2, we also extract the GPS locations of the selected traces for every 20 min to simulate the sensing cycle, then we have 3 time slots of sensing data within the one hour monitoring period. Moreover, we also select 100 traces to sense PM10 signal and 100 traces to sense air humidity signal. By conducting our crowdsensing enhancement approach, we acquire the optimized sensing image (see Figure 16b).
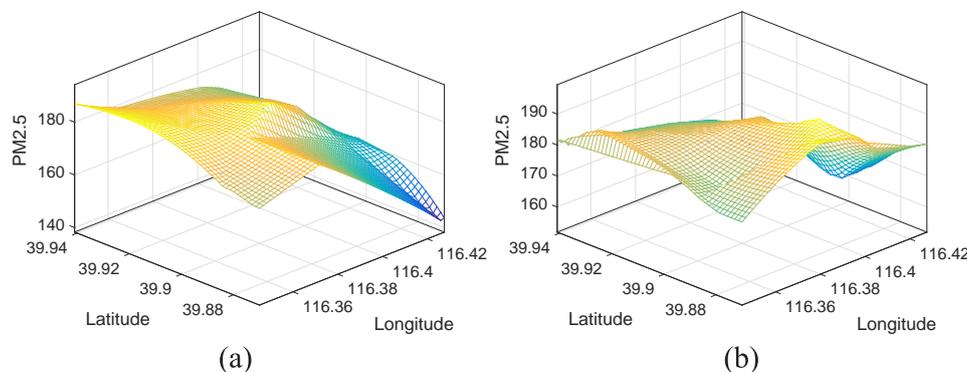
Figure 16. Recovered signals (sensing images). (**a**) The generated sensing image based on interpolation method; (**b**) The generated sensing image based on enhanced crowdsensing approach.

Compared with Figure 16a, the optimized sensing image of Figure 16b holds more details and the correlation coefficient between Figures 15a and 16a is 0.58, while the correlation coefficient between Figures 15a and 16b is 0.91. This result validates the utility of our method.

6. Conclusions

This paper studied the problem of how to promote the sensing ability of crowdsensing network for fine-grained environmental monitoring. In practice, the sensing ability of crowdsensing is limited by the number and space distribution of crowdsensing participants. To enhance the ability of crowdsensing, we proposed a multi-source data driven approach and then utilized a novel metric called "crowdsensing resolution" to quantize the enhancement. The kernel of our enhanced crowdsensing approach, is to leverage temporal and sensory correlated data to help with the recovery of sensing image in blank zones. By improving the sensing ability for blank zones, the crowdsensing network could generate sensing images more precisely, i.e., achieve a higher crowdsensing resolution. More specifically, we built a tensor to model the sensory data and summarized a signal correlation matrix to quantize the correlation between different categories of sensory data. By conducting collaborative tensor decomposition and correlated time slots combination, we presented the optimized sensing images. The numerical simulations and a real dataset based case study verified the resolution promotion of our enhanced crowdsensing beyond traditional interpolation-based approach.

Author Contributions: The work presented here was carried out in collaboration between all authors. Xu Kang, Liang Liu, and Huadong Ma contributed the conception and developed the idea. Xu Kang carried out the experiments and data analysis. Xu Kang and Liang Liu wrote the paper. Huadong Ma provided useful suggestions and helped revise the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yang, X. *Urban Remote Sensing: Monitoring, Synthesis and Modeling in the Urban Environment*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
2. Dong, M.; Ota, K.; Liu, A. RMER: Reliable and Energy Efficient Data Collection for Large-scale Wireless Sensor Networks. *IEEE Internet Things J.* **2016**, *3*, 511–519.
3. Hu, Y.; Dong, M.; Ota, K.; Liu, A. Mobile Target Detection in Wireless Sensor Networks With Adjustable Sensing Frequency. *IEEE Syst. J.* **2016**, *10*, 1160–1171.
4. Liu, X.; Wei, T.; Liu, A. Fast Program Codes Dissemination for Smart Wireless Software Defined Networks. *Sci. Program.* **2016**, *2016*, 1–21.
5. Liu, Y.; Dong, M.; Ota, K.; Liu, A. ActiveTrust: Secure and Trustable Routing in Wireless Sensor Networks. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 2013–2027.
6. Campbell, A.T.; Eisenman, S.B.; Lane, N.D.; Miluzzo, E.; Peterson, R.A.; Lu, H.; Zheng, X.; Musolesi, M.; Fodor, K.; Ahn, G.S. The rise of people-centric sensing. *IEEE Internet Comput.* **2008**, *12*, 12–21.
7. Tang, Z.; Liu, A.; Li, Z.; Choi, Y.J.; Sekiya, H.; Li, J. A Trust-Based Model for Security Cooperating in Vehicular Cloud Computing. *Mobile Inf. Syst.* **2016**, *2016*, 1–22.
8. Ganti, R.K.; Ye, F.; Lei, H. Mobile crowdsensing: Current state and future challenges. *IEEE Commun. Mag.* **2011**, *49*, 32–39.
9. Mitas, L.; Mitasova, H. Spatial interpolation. *Geogr. Inf. Syst. Princ. Tech. Manag. Appl.* **1999**, *1*, 481–492.
10. Liu, L.; Wei, W.; Zhao, D.; Ma, H. Urban resolution: New metric for measuring the quality of urban sensing. *IEEE Trans. Mob. Comput.* **2015**, *14*, 2560–2575.
11. Lee, K.; Hong, S.; Kim, S.J.; Rhee, I. SLAW: Self-Similar Least-Action Human Walk. *IEEE ACM Trans. Netw.* **2012**, *20*, 515–529.
12. Wang, L.; Zhang, D.; Wang, Y.; Chen, C. Sparse mobile crowdsensing: Challenges and opportunities. *IEEE Commun. Mag.* **2016**, *54*, 161–167.
13. Wang, L.; Zhang, D.; Pathak, A.; Chen, C.; Xiong, H.; Yang, D.; Wang, Y. CCS-TA: Quality-guaranteed online task allocation in compressive crowdsensing. In Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing, Osaka, Japan, 7–11 September 2015; pp. 683–694.
14. Zhang, Y.; Roughan, M.; Willinger, W.; Qiu, L. Spatio-temporal compressive sensing and internet traffic matrices. *Acm Sigcomm Comput. Commun. Rev.* **2009**, *39*, 267–278.
15. Li, Z.; Zhu, Y.; Zhu, H.; Li, M. Compressive Sensing Approach to Urban Traffic Sensing. In Proceedings of the International Conference on Distributed Computing Systems, Minneapolis, MN, USA, 20–24 June 2011; pp. 889–898.
16. Rallapalli, S.; Qiu, L.; Zhang, Y.; Chen, Y.C. Exploiting temporal stability and low-rank structure for localization in mobile networks. In Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking, Chicago, IL, USA, 20–24 September 2010; pp. 161–172.
17. Kong, L.; Xia, M.; Liu, X.Y.; Chen, G.; Gu, Y.; Wu, M.Y.; Liu, X. Data Loss and Reconstruction in Wireless Sensor Networks. *IEEE Trans. Parallel Distrib. Syst.* **2014**, *25*, 2818–2828.
18. Xu, L.; Hao, X.; Lane, N.D.; Liu, X.; Moscibroda, T. More with less: Lowering user burden in mobile crowdsourcing through compressive sensing. In Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing, Osaka, Japan, 7–11 September 2015; pp. 659–670.
19. Zhang, Y.; Roughan, M.; Willinger, W.; Qiu, L. Spatio-temporal Compressive Sensing and Internet Traffic Matrices. In Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication, Barcelona, Spain, 16–21 August 2009; pp. 267–278.

20.　Zheng, Y.; Liu, F.; Hsieh, H.P. U-Air: When urban air quality inference meets big data. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, 11–14 August 2013; pp. 1436–1444.

21.　Zheng, Y.; Liu, T.; Wang, Y.; Zhu, Y.; Liu, Y.; Chang, E. Diagnosing New York city's noises with ubiquitous data. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Seattle, WA, USA, 13–17 September 2014; pp. 715–725.

22.　Vardoulakis, S.; Fisher, B.E.; Pericleous, K.; Gonzalez-Flesca, N. Modelling air quality in street canyons: A review. *Atmos. Environ.* **2003**, *37*, 155–182.

23.　Kolda, T.G.; Bader, B.W. Tensor decompositions and applications. *Soc. Ind. Appl. Math. Rev.* **2009**, *51*, 455–500.

24.　Mao, X.; Miao, X.; He, Y.; Li, X.Y. Citysee: Urban CO_2 monitoring with sensors. In Proceedings of the IEEE INFOCOM International Conference on Computer Communications, Orlando, FL, USA, 25–30 March 2012; pp. 1611–1619.

25.　Yuan, J.; Zheng, Y.; Xie, X.; Sun, G. Driving with knowledge from the physical world. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 21–24 August 2011; pp. 316–324.