



Chenyang Zhang¹, Qingli Shi¹, Li Zhuo^{1,2,*}, Fang Wang³ and Haiyan Tao¹

- ¹ Center of Integrated Geographic Information Analaysis, School of Geography Science and Planning, Sun Yat-sen University, Guangzhou 510275, China; zhangchy69@mail2.sysu.edu.cn (C.Z.); shiqli@mail2.sysu.edu.cn (Q.S.); taohy@mail.sysu.edu.cn (H.T.)
- ² Guangdong Key Laboratory for Urbanization and Geo-Simulation, Guangzhou 510275, China
- ³ School of Geographical Science, Guangzhou University, Guangzhou 510006, China; wangfang@gzhu.edu.cn
- * Correspondence: zhuoli@mail.sysu.edu.cn

Abstract: Information on the mixed use of buildings helps understand the status of mixed-use urban vertical land and assists in urban planning decisions. Although a few studies have focused on this topic, the methods they used are quite complex and require manual intervention in extracting different function patterns of buildings, while building recognition rates remain unsatisfying. In this paper, we propose a new method to infer the mixed use of buildings based on a tensor decomposition algorithm, which integrates information from both high-resolution remote sensing images and social sensing data. We selected the Tianhe District of Guangzhou, China to validate our method. The results show that the recognition rate of buildings can reach 98.67%, with an average recognition accuracy of 84%. Our study proves that the tensor decomposition algorithm can extract different function patterns of buildings unsupervised, while remote sensing data can provide key information for inferring building functions. The tensor decomposition-based method can serve as an effective and efficient way to infer the mixed use of buildings, which can achieve better results with simpler steps.

Keywords: mixed-use building; multisource data; tensor decomposition

1. Introduction

Mixed-use buildings refer to buildings that combine multiple functions vertically [1,2]. They represent the vertical dimension of urban land mixing and intensive use [3]. With the development of cities, mixed-use buildings have been growing rapidly for years, and currently exist in various areas of a number of cities [4,5]. The distribution of mixed-use buildings has an important influence on many urban aspects such as traffic, population and the economy. Knowing the distribution of mixed-use buildings can help planners better understand and optimize the status quo of the intensive use of lands, so as to further save travel time and consumption of space and urban energy. Meanwhile, it can also enhance the vitality of buildings and establish a good connection between mixed-use buildings and their surrounding communities and environment, thereby gaining higher economic benefits [6–8]. However, in most areas, the collection of the building function data still relies on time-consuming and laborious manual surveys [9].

Many methods have been proposed to infer building functions, but most of them can only infer the primary function for each building. These methods can be divided into two types according to the data sources used, i.e., remote sensing-based data and social sensingbased data. Research based on remote sensing data usually assumes that there is a certain correlation between building appearances and functions. Various remote sensing images, including high-resolution optical images, stereo optical images and light detection and ranging (LiDAR) data, have been employed to obtain the apparent physical characteristics (e.g., texture, spectral and structural information) of the building, which are further used to infer building functions [9–16]. However, due to the limited correlation between the



Citation: Zhang, C.; Shi, Q.; Zhuo, L.; Wang, F.; Tao, H. Inferring Mixed Use of Buildings with Multisource Data Based on Tensor Decomposition. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 185. https://doi.org/10.3390/ijgi10030185

Academic Editors: Giuseppe Borruso and Wolfgang Kainz

Received: 4 February 2021 Accepted: 18 March 2021 Published: 20 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). apparent physical characteristics of a building and its function, the appearance cannot fully infer the functions of all building types. Meanwhile, research based on social sensing data (such as mobile phone data, social media data, taxi trajectory data and point of interest (POI) data) assume that buildings with the same functions have similar human activities [17]. They focus on extracting characteristics of human activities from social sensing data and then use clustering-based methods or other machine learning methods to infer a building's single-use function [18–22]. Compared with the remote sensing-based methods, these methods can obtain more detailed building functions [23,24]. However, the coverage of social sensing data may also affect their results. For example, taxi pick-up and drop point data are only distributed on the roads, so human activity characteristics extracted from such data would therefore require certain ways to link to the adjacent buildings.

Few studies have focused on inferring a building's mixed-use functions, most of which are based on social sensing data. Niu et al. [25] first proposed a density-based method to characterize mixed-use buildings on the basis of the assumption that buildings with similar functions would have similar peak times in terms of taxi passenger pick-ups or drop-offs and Tencent user activity. Liu et al. [26] improved Niu et al.'s method by inferring a building's mixed-use functions on the basis of the purpose of the trips obtained from the taxi data. These research studies, however, still have limitations. First, in such research, characteristics like the "peak time" of each type of building are extracted on the basis of training samples, which is a supervised classification method and relies more on experience. Due to the lack of automation, using these supervised classification methods to extract human activity characteristics is quite complicated. Second, the "peak time" of human activities extracted directly from mixed-use building temporal human activity curves is unstable. The temporal human activity curve of a mixed-use buildings is the mixing and superimposing result of temporal human activity curves that belong to different crowds with different activity patterns. The characteristics of such a temporal human activity curve may vary with the population groups attracted by the building's different functions. Finally, the building recognition rates of these studies are not satisfying. This is probably because they used only social sensing data, which has shortages like low resolution and coverage.

To overcome the limitations of existing studies, a tensor decomposition-based method that integrates information from both high-resolution remote sensing images and social sensing data is proposed in this paper to infer a building's mixed-use functions. As a high-order generalization of the matrix singular value decomposition and principal component analysis, tensor decomposition not only can decompose the n-dimensional array to obtain the different characteristic patterns as well as the correlation between objects and characteristic patterns in each dimension, but also can obtain the relationship between characteristic patterns in different dimensions. Considering that the temporal population distribution inside buildings has space-time dimensions, by constructing a tensor of changes in human activities within the building over time and decomposing it, we can obtain different human activity patterns inside buildings unsupervised and infer the mixed use of the building through these patterns. Therefore, tensor decomposition is very suitable for the inference of a building's mixed-use functions. Two advantages of our study are as follows: First, we realized the unsupervised and automatic extraction of the stable human activity characteristics inside buildings by applying a tensor decomposition algorithm. Second, we improved the building recognition rate by integrating high-resolution remote sensing images and social sensing data. A case study in Tianhe District, Guangzhou, China illustrates the advantages of our method. The rest of this paper is organized as follows. The study area and datasets are described in Section 2. Section 3 provides a detailed description of the tensor decomposition-based method, and Sections 4 and 5 report the experimental results and discussion, respectively. Finally, the study is concluded in Section 6.

2. Study Area and Datasets

As shown in Figure 1, in this study, we selected Tianhe District of Guangzhou in China as our study area. Located in the eastern part of Guangzhou, it had a population of 1,545,700

in 2015 and a total area of 137.38 square kilometers (Bureau of Statistics of Guangzhou 2015, http://112.94.72.17/portal/queryInfo/statisticsYearbook/index, accessed on 19 August 2020). As the new downtown area and commercial center of Guangzhou, Tianhe District has a large number of commercial and office buildings. Meanwhile, there are also many industrial buildings and urban villages in the Tianhe District. The variety of building types and the highly mixed building functions make Tianhe District an ideal study area for this study.



Figure 1. Location of the Tianhe District and its building footprints.

Three different datasets, including Tencent user density data, Worldview2 highresolution images and building footprints data, were used in our research.

The Tencent user density data record the number of smartphone users accessing Tencent's real-time location service products per hour [26]. Thanks to the enormous user amount, it could serve as an ideal indicator to represent real-time human activities. In this study, we collected the Tencent user density data from the Easygo platform, which covered the period from 15 June to 21 June 2015. After format conversion and the data cleaning pre-process, we got the Tencent user density data in the format of sample points with a spatial resolution of 25 m and a temporal resolution of one hour.

The building footprints data were obtained from Baidu map platform and had 23,446 building footprints in the study area. Each building footprint represents a single building that was used as one of the basic units for the building mixed-use inference.

Worldview2 images employed in this study were recorded on November 14, 2010 (Processing Level: Ortho Ready Standard) and can be found on the website of the provider (http://worldview2.digitalglobe.com/, accessed on 23 March 2014). The 1.8 m spatial resolution image had four standard spectral bands, namely, the blue band (0.45–0.51 μ m), the green band (0.51–0.58 μ m), the red band (0.63–0.69 μ m) and the near infrared band (0.77–0.90 μ m).

Although there were differences in the acquisition time for these three datasets, it was still acceptable given the fact that most buildings remained unchanged during the period when the above-mentioned data were recorded. The coordinates of all three datasets were converted to the Zone 49 UTM (Universal Transverse Mercator) projected coordinate system.

3. Method

In this study, we developed a tensor decomposition-based method to infer buildings' mixed-use functions. Figure 2 shows the flowchart of the proposed method, which can be divided into three steps. First, the adjacent single buildings with similar appearance were merged into building groups based on features extracted from Worldview2 high-resolution images. Then, based on human activity characteristics extracted from Tencent user density data, a building dynamic characteristic tensor was constructed and further decomposed. Finally, buildings' mixed-use functions were inferred according to the tensor decomposition results. Details of each step are described below.



Figure 2. Flowchart of inferring buildings' mixed-use functions.

3.1. Single Building Merging

The spatial resolution of the Tencent user density data was 25 m, which means that not every single building can be covered by this data, as shown in Figure 3a. Moreover, most small single buildings only have one sampling point of Tencent user data, which is not stable enough to extract their temporal human activity curves. Residential buildings located in the same residential area or working buildings in a factory often have a highly consistent appearance and spatial adjacent. We think that single buildings with adjacent locations and similar appearances are very likely to have the same function. Therefore, we merged adjacent single buildings with similar appearance into building groups, and used building groups as the objects to identify the building's mixed-use functions based on social sensing data. In this way, we solved the above-mentioned problems by converting the study unit from a single building to a building group.

Worldview2 high-resolution images were used to merge single buildings. We first used the building footprint vector data to clip Worldview2 high-resolution images to obtain the high-resolution remote sensing image for each single building. Since the building footprint vector data matched well with the Worldview2 image (see Figure 3a), we did not need to register two kinds of data. Then, we assigned a unique ID number to each single building, and used the eCognition software to extract the apparent characteristics from the respective Worldview2 images. The eCognition software is a remote sensing image processing software. We import the cropped Worldview2 images into this software to generate an object layer. The Worldview2 image of each single building was regarded as an object. For each object in this object layer, eCognition software can provide three types of feature extraction algorithms: spectral characteristics, shape characteristics and texture characteristics. The user only needs to select the required characteristics, and the software can automatically calculate these characteristics of all the objects in the object layer. We choose to extract the spectral characteristics, such as the mean, standard deviation of each band and the shape characteristics, such as the area, perimeter and shape index. Finally, the apparent characteristics of single buildings were added to the attribute table of the building footprint vector data.



Figure 3. (**a**) Building footprints and Tencent user data in a sample area in Tianhe District, and (**b**) Building footprints and Worldview2 image in the same area.

The group analysis tool in ArcMap was used to merge single buildings into building groups on the basis of the building footprint vector data, which had the attribute table of apparent characteristics. This tool is a packaged tool in the ArcMap toolbox. It performs a classification process to find natural clusters in the data. After specifying the number of groups, it looks for a solution that makes all the elements in each group as similar as possible, but as different as possible among groups. The element similarity is measured on the basis of a set of characteristics specified for the objects to be analyzed. We only need to enter the vector data of single buildings and specify the building apparent characteristic attributes upon which the grouping is based and the number of groups, and the tool will automatically generate group analysis results and assign respective group IDs to all the single buildings. Single buildings assigned the same group ID were a building group. After grouping all the single buildings, the building group referred to a group of single buildings with the same appearance or a single building with appearance obviously different from its surroundings. To make sure that the number of groups was reasonable, we divided the entire study area into many smaller areas and performed group analysis in these different areas. At the same time, we assigned as many groups as possible to each small area to ensure that some single buildings with highly similar appearances and adjacent spaces were grouped into one group. These building groups were then used to replace single buildings in the subsequent steps of the building mixed-use inference.

3.2. Construction and Decomposition of the Building Dynamic Characteristic Tensor

An n-dimension array is defined as a tensor. Tensor decomposition is a higher-order generalization of the matrix singular value decomposition and principal component analysis, which is commonly used to eliminate the correlation among features in vector spaces and perform feature selections of spatio-temporal data [27]. Tucker decomposition is one of the main tensor decomposition methods. It not only can decompose the n-dimensional array to obtain different characteristic patterns in each dimension, but also can obtain the relationship between characteristic patterns in different dimensions [27,28]. The Tencent user density data can reflect changes of population distribution within buildings over time, which means it has both space and time dimensions. We can, therefore, construct a tensor that reflects the changes of human activities inside buildings over time. By decomposing

this tensor, we can obtain different human activity patterns inside the building from the spatial dimension, as well as time characteristics of different human activity patterns, which can be further used to determine the purpose of human activities in this pattern. Since studies have found that the purpose of human activity inside a building is closely related to the function of the building [26], it is therefore possible to further infer the mixed use of buildings based on the above-mentioned decomposition results.

The Tencent user density data records the number of Tencent users per hour at each sampling point with a sampling spacing of 25 m. It can reflect the human activity changes over time. In this study, we counted the hourly average value of the Tencent user density sample points inside the building groups to characterize their vitality. High user density indicates great vitality. The spatio-temporal distribution of the 1-week vitality of the *N* building groups in the study area, B_D , can be characterized as follows.

$$B_D = \begin{pmatrix} B_1^1 & \cdots & B_1^{168} \\ \vdots & \ddots & \vdots \\ B_N^1 & \cdots & B_N^{168} \end{pmatrix}$$
(1)

where B_n^t represents the Tencent user density value of the building group *n* at hour *t* and *n* is belongs to [1, *N*]. *N* is total number of building groups.

Furthermore, the three-order tensor T responding to the building vitality can be built based on B_D . As shown in Figure 4, the three dimensions of T represent the buildings, time and date, respectively. Each frontal slice of the tensor T represents a day, each row of the frontal slice represents a building group and each column represents an hour of a day. The value of a grid element represents the Tencent user density of a building group within one hour of a day.



Figure 4. Tucker decomposition of a three-way array tensor *T*.

Tucker decomposition decomposes a tensor into a core tensor multiplied (or transformed) by a matrix along each mode [27]. Taking the third-order tensor $\chi \in R^{(I \times J \times K)}$ as an example, a Tucker decomposition of χ with the number of decomposition components of three dimensions (set as a, b and c) yields a core matrix $G \in R^{(a \times b \times c)}$ and factor matrices $A \in R^{(I \times a)}$, $B \in R^{(J \times b)}$ and $C \in R^{(K \times c)}$ on the three dimensions. Each column of the factor matrix is a component representing a pattern in the feature space, while each row represents values of an object in different patterns, indicating the correlation between that object and different patterns. A larger pattern value indicates a closer correlation between the object and the pattern. The core tensor represents the intensity of the interactions between patterns in different factor matrices. Based on these theories, we can decompose the tensor *T* into the three factor matrices *A*, *B* and *C* and a core tensor *G*. Factor matrices *A*, *B* and *C* represent the building function mode, the time mode and the date mode, respectively, while the core tensor *G* reflects the relationship between the three factor matrices' patterns, as shown in Figure 3 and Equation (2). In function factor matrix A ($N \times k$), each

row represents a building group and each column represents a building functional pattern, while the pattern values in matrix *A* represent the correlation between the building groups and building functional patterns. In time factor matrix *B*, each row represents an hour of the day and each column represents a time pattern, while the pattern values in matrix *B* represent the correlation between the hours of the day and time patterns. In date factor matrix *C*, each row represents a day of the week and each column represents a date pattern, while the pattern values of matrix *C* represent the correlation of days of the week with date patterns. After determining the building types of *k* functional patterns, we can infer the building's mixed-use functions based on matrix *A*.

$$T \approx G_{k \times 4 \times 2} \times A_{N \times k} \times B_{24 \times 4} \times C_{7 \times 2} \tag{2}$$

According to the existing research [29], we set the pattern number of the time factor matrix *B* as 4, as shown in Figure 3, representing the morning, afternoon, evening and night patterns, respectively; we set the pattern number of the date factor matrix *C* as 2, representing the weekday pattern and weekend pattern, respectively. The k patterns of the factor matrix *A* represent the different human activity patterns. Since human activities inside certain building are closely related to the building's function, the *k* patterns of the factor matrix *A* can thereby be regarded as the different building functions. The value of *k* was determined by minimizing the cost function in the tensor decomposition process. After determining the building types of *k* functional patterns, we can infer the building's mixed-use functions based on matrix *A*.

In the process of Tucker decomposition, we adopted the sparse non-negative Tucker decomposition (SNTUCKER) method proposed by Mørup et al. [30]. This method incorporates both sparse and non-negative constraints in the Tucker decomposition, which ensures the non-negativity of the factor matrix and makes the results as sparse as possible to obtain the main features, thus reducing the ambiguity of the decomposition results.

3.3. Building Mixed-Use Extrapolation

3.3.1. Pattern Inference for Factor Matrices

Since the tensor decomposition is a blind decomposition process, it is necessary to link the building functions with the k patterns of the functional factor matrix A. The specific steps are as follows. First, find the correlation between each time pattern in the time factor matrix B and the 24 h in one day and infer the time patterns (morning, afternoon, evening and night modes) corresponding to the four patterns in matrix B. Second, find the correlation between each principal pattern in the date factor matrix C and each day of the week, and infer the day patterns corresponding to the two patterns in matrix C (weekday pattern and weekend pattern). Finally, find the time and date patterns that are most relevant to the k functional patterns.

3.3.2. Building Mixed-Use Inference

After determining the building function represented by each pattern of matrix A, we can further infer the building function of each building group. As each row of the functional factor matrix A represents the correlation between a building group and k functional patterns. For each building group, the functional pattern corresponding to the largest pattern value represents its main function. Considering there were also many mixed-use buildings, we set the degree of functional mixing (D_i) to determine whether a building is a mixed-use building, and to identify its mixed-use functions, as shown in Equation (3).

$$D_i = 1 - \frac{M1_i - M2_i}{M1_i}$$
(3)

where $M1_i$ represents the largest functional pattern value of the *i*th building group, i.e., the largest pattern value of the *i*th row in the functional factor matrix A, while $M2_i$ represents

the second largest functional pattern value of the *i*th building group, i.e., the second largest pattern value of the *i*th row in the functional factor matrix A. In Equation (3), the difference between $M1_i$ and $M2_i$ indicates the similarity of $M2_i$ and $M1_i$. We believe that if the difference between $M1_i$ and $M2_i$ is too large, the building group can only significantly reflect a function represented by $M1_i$. If the values of $M2_i$ and $M1_i$ are similar, the two functions represented by $M2_i$ and $M1_i$ are both owned by the building group. Here, we only considered the case where a building had at most two functions. To make the values of $(M1_i - M2_i)$ comparable between the different building groups, we used $(M1_i - M2_i)/M1_i$ to normalize them. It can be seen that the smaller $(M1_i - M2_i)/M1_i$ is, the more likely that the building group has mixed-use functions. Thus, a greater D_i value indicates a higher degree of mixed building functions. When D_i exceeded a certain threshold, we defined the building as a mixed-use building. To determine the threshold of D_i , we selected samples of each type of single-use building and mixed-use building in the study area and calculated their mixing degree, D_i , respectively. The number of samples of each category is shown in the first column of Table 1. It can be seen that all types of single-use buildings had an average value of D_i less than 0.7, while the average D_i values of the mixed-use buildings types were greater than 0.7. Therefore, we set the threshold of D_i as 0.7 in this study. That is, when the D_i of the building group was greater than or equal to 0.7, the building group was regarded as a mixed-use building, whose functions correspond to its maximum and second-largest functional pattern value. Otherwise, the building group was determined to be a single-use building, with its function corresponding to the maximum functional pattern value.

Table 1. Statistics of the Average *D_i* Values of Each Type of Building Group Sample.

Type and Number of Samples	Largest Pattern Value	Second-Largest Pattern Value	Average D _i
Working (87)	0.0077	0.0045	0.5844
Residential (153)	0.0081	0.0034	0.4198
Shopping and recreational (19)	0.0051	0.0035	0.6667
Residential and working (31)	0.0021	0.0017	0.7941
Working and recreational (22)	0.0040	0.0030	0.7500
Residential and recreational (24)	0.0033	0.0025	0.7567

3.4. Accuracy Evaluation

Without real building mixed-use data for the study area, this research selected some sample areas to evaluate the building mixed-use recognition results. We first used the recognition rate to represent the proportion of buildings that can be recognized by this method in all buildings, as shown in Equation (4). We then referred to the accuracy evaluation method of remote sensing image classification and we constructed the confusion matrix for each sample area and calculated the overall accuracy (OA), the Kappa coefficient, the sensitivity and the precision to test the classification accuracy.

$$RT = \frac{Num_r}{Num_{all}} \tag{4}$$

where *RT* is the recognition rate, Num_r is the number of single buildings that are assigned categories in the sample areas and Num_{all} is the total number of single buildings in the sample areas. The recognition rate can reflect whether our method can effectively reduce the number of unidentified buildings compared to the existing methods.

The selection method of the sample area is as follows. We first divided the study area into grids with resolutions of 500 m \times 500 m and 1 km \times 1 km, respectively, and then randomly selected grid samples at both spatial scales to calculate the recognition rate and accuracy rate of the building functions.

4. Results

4.1. Results of Single Building Merging

After merging single buildings, we obtained a total of 8750 building groups. Figure 5 shows the building groups and the corresponding Worldview2 images of the three sample areas in the Tianhe District, which reflect the different types of building groups. Figure 5b shows the single building merging result of commercial and residential areas. By comparing this result with its corresponding Worldview2 image in Figure 5a, it can be seen that the two large commercial and entertainment buildings were not merged with other single buildings because their shapes and apparent characteristics are quite different from those of the surrounding single buildings. Instead, they were directly given an independent ID as a separate building group. Figure 5d reflects the single building merging result in a sample area of office buildings. Comparing it with Figure 5c, it can be seen that office buildings with a large area and special appearance were not merged, while those with a small area and consistent appearance were merged into a group. Figure 5f shows the single building merging results of factory areas and urban villages. Compared with Figure 5e, it can be seen that the urban villages have relatively special spectral and shape characteristics. Because of the small building area and dense distribution, the urban village area in the building footprint data was drawn as continuous irregular rectangular areas. On the one hand, the building groups of urban village buildings were larger than other types of building groups, thanks to their large area and strong appearance consistency. On the other hand, many small footprints at the edge of these areas were not merged because of their irregular shapes. To sum up, the single building merging results were quite reasonable. For example, most of the commercial and entertainment single buildings and larger office single buildings were not merged, and thus, can be directly used as building groups to participate in building function inference. On the other hand, the relatively small residential single buildings, office single buildings, urban village buildings and industrial single buildings were usually merged because of the highly consistent appearance of the buildings around them. This solved the defect that many single buildings with small areas cannot be covered by Tencent user data. Meanwhile, it also ensured the independence of buildings with large areas, such as commercial entertainment and office buildings, which could better meet the requirements of the subsequent building mixed-use inference.



Figure 5. Single building merging results, (**b**,**d**,**f**), and their corresponding Worldview2 images, (**a**,**c**,**e**), of the three sample areas in the Tianhe District.

4.2. Analysis of Tensor Decomposition Results

Before constructing the tensor *T* that responds to the building vitality, we determined the pattern number *k* of the functional factor matrix *A* by minimizing the cost function. We conducted five sets of tensor decomposition experiments, with different *k* values in the range of 3–15. As shown in Figure 6, the five curves represent the five groups of experiments, respectively. It can be seen that for all the five groups of experiments, the cost function curves all decreased rapidly when *k* increased from 1 to 7, but became quite stable after that. Therefore we set the value of *k* to 8 and the size of the core tensor *G* as $8 \times 4 \times 2$.



Figure 6. The change of objective function with different k values.

4.2.1. Time Pattern Analysis

Figure 7a,b shows the decomposition results of the time factor matrix and the date factor matrix, respectively. The pattern values in the factor matrix were normalized by the column for the purpose of comparison. A larger pattern value indicates a closer correlation between the object and the pattern. Figure 7a reflects the relationship between the 24 h of a day and the four time patterns. The pattern value of Pattern 1 began to grow at 10:00 a.m., reaching its maximum at 8:00 p.m., thus making Pattern 1 the night pattern. By contrast, the pattern value of Pattern 2 increased significantly at 8:00 a.m., reaching the maximum at 11:00 a.m., making it more fit for the morning pattern. With larger pattern values during the night period, Pattern 3 can also be classified as the night pattern. Pattern 4's pattern value started to increase from 10:00 a.m. and peaked at 6:00 p.m., proving that it is an afternoon pattern. Similarly, Figure 7b shows the correlation between each of the days from 15 June to 21 June 2015 and the two date patterns. The 20 and 21 June 2015 were weekends, while the other dates were weekdays. It was clear that Pattern 1 had greater values on weekdays, but smaller values on the weekends, whereas Pattern 2 showed quite the opposite trends. This proves that Pattern 1 represents the weekday pattern, whereas Pattern 2 represents the weekend pattern.

4.2.2. Building Functional Pattern Inference

The core tensor of *T* reflects the connection between different patterns of the functional factor matrix A, the time factor matrix *B* and the date factor matrix C. Since building function can be determined according to the characteristics of the internal human activities, the building functions of eight function patterns can be inferred by analyzing the correspondence between the time, date patterns and the building function patterns through the core matrix.



Figure 7. Factor matrix characteristic patterns (**a**) correlation of time patterns with 24 h; (**b**) correlation of date patterns with 7 days.

Table 2 shows the core tensor obtained from the tensor T decomposition. The values in the table correspond to the links' strength between the eight functional patterns, the four time patterns and the two date patterns. Values with stronger ties are highlighted in orange color in Table 2.

	Weekday			Weekend				
	Morning	Afternoon	Evening	Night	Morning	Afternoon	Evening	Night
Pattern 1	0.003	0.003	0	0.156 ¹	0.093	0.085	0.030	0.261 ¹
Pattern 2	0	0	0	0	0.626 ¹	0.007	0	0.008
Pattern 3	0	0.002	0.052	0	0	0.271 ¹	0.235 ¹	0.012
Pattern 4	0	0.259 ¹	0	0	0	0.022	0	0.028
Pattern 5	0.061	0.002	0.113 ¹	0	0	0	0.311 ¹	0.223 ¹
Pattern 6	0.552 ¹	0.006	0.007	0	0.074	0.003	0	0.001
Pattern 7	0.016	0.244 ¹	0.349 ¹	0	0	0.053	0.001	0.008
Pattern 8	0.005	0	0.099	0.365	0	0.004	0.004	0.052

Table 2. Core Tensor of	Τ.
-------------------------	----

¹ The highlight in the table indicates the value that reflects the larger correlation relationship between different patterns.

Based on results in Table 2, the eight building functional patterns are summarized as Table 3. Pattern 1 and Pattern 8 in the function factor matrix *A* were strong correlated with the night pattern. Since buildings with the night pattern have peak times of human activities at night, we can infer that Pattern 1 and Pattern 8 correspond to the living function. Pattern 5 in function factor matrix *A* was correlated with three patterns, namely, the evening pattern on weekdays, and both the evening and night patterns on weekends, which were thus determined as the living function. Patterns 4, 6 and 7 in the function factor matrix *A* were only related to the weekday pattern. Patterns 4 and 6 corresponded to the afternoon pattern and the morning pattern. Despite the slight difference in the peak times of human activities with these three patterns, they all reflected the characteristics of work-related

building functions. In contrast to Patterns 4, 6 and 7, Patterns 2 and 3 both showed a closer relation to the daytime pattern on the weekend. Pattern 2 corresponded to the morning pattern on weekends, reflecting the characteristics of buildings in the market, while Pattern 3 was more closely related to the afternoon and evening patterns on weekends, reflecting the typical characteristics of buildings such as shopping malls. In this study, Patterns 2 and 3 were combined as shopping and recreational functions.

	Weekday	Weekend
Morning	Pattern 6	Pattern 2
Afternoon	Pattern 4,7	Pattern 3
Evening	Pattern 7,5	Pattern 3,5
Night	Pattern 1,8	Pattern 1,5

Table 3. Functional Pattern and Temporal Pattern Interaction.

To further validate the inferred results, we plotted the curves of the average number of Tencent users in all buildings corresponding to the eight types of building functions over time, and observed the human activity patterns in buildings with different functions. Here, we did not consider the mixed use of the building. The function of each building group was determined according to the maximum function mode value corresponding to the function factor matrix A. We separately counted the average weekdays and weekends of the eight different functional buildings with temporal human activity curves, as shown in Figure 8. The values of 1–24 on the horizontal axis represent the averaged 24 h of the weekdays, while the values of 25–48 represent the averaged 24 h of the weekends. The vertical axis represents the average number of Tencent users of a specific type of building at a certain time, which reflects the population activities inside those buildings at that time. It can be observed that all three curves reflect the distinct characteristics of residential buildings, i.e., showing the peak times of human activity around 11:00 p.m. Pattern 1 and Pattern 5 maintained the same population on weekdays and weekends, while it was different in Pattern 8, which showed a significantly smaller number of people on weekends than weekdays. This indicates that there is an outflow of the population from buildings of Pattern 8 during the weekend. Figure 8b shows the average temporal human activity curves corresponding to the three working functional patterns. All three curves showed clear peak times of human activity in the morning and afternoon on weekdays, becoming smooth on the weekends. These three curves, compared with the weekdays, showed a significant decrease in the number of people on the weekends, which are typical curves of the working function. The difference among the three curves was that the afternoon human activities of Pattern 4 were significantly more active than the morning human activities on the weekdays; Pattern 7 showed similar human activities in the morning and afternoon; and Pattern 6 demonstrated significantly more active human activities in the morning than in the afternoon. This may be related to the different natures of the work of people within office buildings. Figure 8c shows the average temporal human activity curves corresponding to the two functional patterns classified as the shopping and recreation functions. Both curves were characterized by a greater weekend pedestrian volume than weekday pedestrian volume. The peak times of the human activities of Pattern 2 occurred at 10:00 am, conforming to the characteristics of the densest crowd in the trading market in the morning, while the curve of Pattern 3 displayed the peak times of human activities at 5:00 p.m., conforming to the characteristics of shopping and recreation buildings. In summary, the function patterns reflected in average population density characteristic curves were fully consistent with the inferred results based on the core tensor, thus proving the reasonableness of the inference obtained with the core tensor.



Figure 8. Average temporal human activity curves of eight function patterns: (**a**) characteristic curves corresponding to the residential function, (**b**) characteristic curves corresponding to the working function, and (**c**) characteristic curves corresponding to the shopping and recreational function.

4.2.3. Building Function Recognition Accuracy Assessment

The inference results of building mixed-use functions in the Tianhe District are shown in Figure 9. Figure 9a shows the distribution of different types of buildings. Overall, the number of residential buildings in the Tianhe District was significantly higher than other types of buildings. They were evenly distributed throughout the Tianhe District, except for the south-west part. The working buildings showed a relatively clustered distribution, with the densest distribution located in the Southwest Tianhe District, i.e., the central business district (CBD) of Guangzhou, known as Tianhe City and Zhujiang New Town. Figure 9b shows the enlarged map of Zhujiang New Town. In other areas of the Tianhe District, there were a number of relatively small clusters of working buildings, most of which can be associated with industrial parks. Shopping and recreational buildings were fewer in number and spatially discrete, most of which were located in the southern part of the Tianhe District. Similarly, the density of the shopping and recreational buildings in the CBD area was also higher than other areas.



Figure 9. Buildings' mixed-use inference results in the Tianhe District: (**a**) building function inference results, (**b**,**c**) local enlarged views, and (**d**) distribution of mixed-use buildings.

Figure 9d shows the distribution of mixed-use buildings in the Tianhe District. It can be seen that the mixed-use buildings in the Tianhe District were mainly distributed along the roads. Again, the number of mixed-use buildings in the southwest part of the Tianhe District was significantly higher than in other areas.

Based on our results, we further calculated the proportion of each building type in the study area (Table 4). The values of the "Number" column were obtained by counting the numbers of each type of single building, while the values of the "Percentage" column were calculated by dividing the number of each type of single building with the total number of single buildings in the Tianhe District. From Table 5, it can be seen that the number of mixed-use buildings in the Tianhe District accounted for 18.92% of all buildings, which is consistent with the results of existing studies [26]. Among the mixed-use buildings, working and residential buildings accounted for the largest proportion, i.e., 11.10% of the total number of buildings for 4.55% of the total number of buildings. The work and recreation buildings group only accounted for 3.27%. The buildings' recognition rate in our method reached 98.67%, which is a significant improvement compared to the existing studies on building function inference [20,25].

Туре	Number	Percentage (%)
Single-use building	18,219	79.75
Mixed-use building	4322	18.92
Working building	2380	10.42
Residential building	14,715	64.41
Shopping and recreational building	1124	4.92
Residential and working building	2536	11.10
Working and recreational building	747	3.27
Residential and recreational building	1039	4.55
Unclassified building	304	1.33

Table 4. Number and Percentage of All Types of Single Buildings in the Tianhe District.

To verify the accuracy of the building function recognition, we randomly selected six 500 m \times 500 m and 1000 m \times 1000 m grids in the Tianhe District as the sample areas. The actual function types of the single buildings in the sample areas were marked on the basis of both a field survey and the Baidu street view map, which were then used to validate the

recognition results in our study. Table 5 shows the OA, Kappa coefficient and recognition rate of the mixed use of the single buildings in two spatial scales. The results show that the proposed method achieved a high accuracy of building mixed-use recognition results with an average OA of 0.84 and an average Kappa of 0.75, which prove that our method can effectively infer the mixed-use functions of buildings. To further explore the difference in the recognition accuracy of different types of buildings, we constructed a confusion matrix with a total of 716 buildings in 6 sample areas and calculated the sensitivity and precision of each building type, as shown in Table 6. As can be derived from Table 6, the recognition accuracy of single-use buildings was higher than that of mixed-use buildings. Among the types of single-use building, working buildings had the highest recognition accuracy, with a sensitivity of 95% and a precision of 96%, followed by residential buildings, and finally, shopping and recreational buildings. As for the mixed-use buildings, the identification of working and shopping building was the most accurate, with a sensitivity of 74% and a precision of 95%, and the identification of residential and shopping building was the least accurate. This is mainly because the low-level shops in these residential and shopping buildings cannot attract enough people, so their commercial characteristic curves are not recognized in the recognition process. Therefore, these mixed-use buildings are mistakenly classified as residential buildings.

Spatial Scale	Samples	Number of Single Buildings	Number of Correctly Inferred Single Buildings	Recognition Rate (RT)	OA	Kappa
500 m × 500 m		76	67	1	0.88	0.75
		23	19	1	0.83	0.75
		109	89	1	0.82	0.64
		216	181	1	0.84	0.75
$1 \text{ km} \times 1 \text{ km}$		86	71	1	0.83	0.77
		206	171	1	0.83	0.73
Total		716	598	1	0.84	0.75

Table 5. Accuracy Assessment in Different Spatial Scale.

Table 6. The sensitivity and precision of building mixed-use inference results in six sample areas.

Туре	Sensitivity	Precision
Working building	0.95	0.96
Residential building	0.90	0.94
Shopping and recreational building	0.54	0.77
Residential and working building	0.67	0.64
Working and shopping building	0.74	0.95
Residential and shopping building	0.80	0.45

5. Discussion

5.1. Error Analysis of Building Function Inference

The accuracy of building function inference results based on the tensor decomposition was different among the different building types. In addition to the mixed-use buildings analyzed in Table 5, where only one building function was identified and mistakenly classified as a single-use building, there were also cases where one type of single-use building was clearly mistaken as another type of single-function building. We selected the typical misclassified buildings in Zhujiang New Town and Tianhe City, which have the highest complexity of building types in Tianhe District, and analyzed the main reasons for the classification errors.

Table 7 shows the major misclassification conditions of each type of building. Buildings which were misclassified as residential buildings were mainly sports venues and schools, while some of the high-end residential areas were mistakenly classified as working buildings. The obvious misclassification of shopping and recreational buildings only existed in some of the uncompleted or completed villa areas.

Misclassified Building Function	Examples
Residential building	Guangzhou Tianhe Stadium, Guangzhou Tianhe Sports Center, Haixinsha Asian Games Stand, Tianhe Huayang Primary School (Huacheng Campus)
Working building	Poly 108 Mansion, Bihai Bay Residence, Jiayu Junyue Residence
Shopping and recreational building	Hesheng DiJing Villa, Longfor Mansion villa

Table 7. Typical Misclassified Buildings in Zhujiang New Town and Tianhe City.

We concluded the three main reasons that led to the incorrect inference of building functions from Table 6. Firstly, inference of building functions mainly depends on the activity patterns of the people inside buildings. If the characteristics of human activities in buildings cannot be accurately reflected, the building function inference results will be biased. For example, the reason for misclassified sports venues is that their temporal human activity curves have similar characteristics to the temporal human activity curves of residential buildings, because people tend to go to sports venues for exercise in the evening. Secondly, the representativeness and density of social sensing data affects the inference results. The human activities represented by Tencent user density data more often reflect the activities of groups of adults who frequently use mobile devices, thus affecting the accuracy of the results that are mainly associated with the elderly and student groups. The classic incorrect inference caused by this reason relates to school buildings. Finally, when we merged the functional patterns obtained from the tensor decomposition into three functional patterns, we did not further explore the intra-class differences among the three functional patterns, which may have omitted certain information and affected the accuracy of the inference of building functions. Given that these intra-class differences can reflect the building's functions to a certain extent, it is necessary to further explore the decomposition results in conjunction with other auxiliary data in order to obtain more detailed and accurate building function inference results.

5.2. Advantages of Integrating Remote Sensing Data to Infer Building Mixed-Use

In this study, we introduced high-resolution remote sensing data to merge buildings with similar appearance into building groups. There were two main reasons for doing so. First, the 25 m resolution of the Tencent user density data was not sufficient enough to cover all the buildings. Creating buffers for single buildings was not viable either, because the human activity characteristics of the adjacent buildings would influence each other, especially in areas with a high complexity of building types. Second, the human activity characteristics of some buildings could not be fully represented due to the limited number

of Tencent user density data sampling points. Since building groups are larger than single buildings in size, the coverage of Tencent user density data could be greatly increased, which can effectively improve the recognition rate of building functions. The percentage of unidentified buildings was only 1.33% when using the building groups, in contrast to 32.39% when using the single buildings directly. This demonstrates the necessity of merging single buildings on the basis of high-resolution images. Moreover, the human activity characteristics extracted on the basis of the building groups were more stable than those extracted based on single buildings, thus allowing us to obtain more accurate building mixed-use inference results.

6. Conclusions

This study proposed a new method to infer buildings' mixed-use function, which was based on the integrative use of high-resolution remote sensing images and social sensing data, as well as the tensor decomposition algorithm. We first extracted building apparent physical characteristics from the high-resolution remote sensing data and merged single buildings with similar appearance into building groups. Then, we used the Tencent user density data to construct building dynamic characteristic tensors for the building groups. Finally, we inferred the buildings' mixed-use functions based on the tensor decomposition results. The application of the proposed method in the Tianhe District of Guangzhou, China yielded a building recognition rate of 98.67%, with an average recognition accuracy of 84%. The following conclusions can be made based on the results of this study. First, the integration of high-resolution remote sensing images and social sensing data helped to increase the building function recognition rate from 67.61% to 98.67%, compared to the case of using social sensing data only. This indicates that combining these two types of data can effectively improve the building recognition rate. Second, the building function inference method based on tensor decomposition had an average accuracy of 84%, which proves that the tensor decomposition algorithm can accurately identify different function patterns of buildings. However, the proposed method can only extract three relatively rough functional categories, namely residence, work and business entertainment. Future studies may consider adding other data, such as POI, to obtain betterbuilding function inference results.

Author Contributions: Conceptualization, Chenyang Zhang, Li Zhuo and Qingli Shi; Funding acquisition, Li Zhuo and Fang Wang; Methodology, Qingli Shi and Chengyang Zhang; Supervision, Li Zhuo, Fang Wang and Haiyan Tao; Writing—Original Draft Preparation, Chenyang Zhang; Writing—Review and Editing, Li Zho, Fang Wang and Haiyan Tao. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Science Foundation of China (Grant No. 41971372, Grant No. 41971345, and Grant No. 42071262), in part by the Natural Science Foundation of Guangdong Province (Grant No. 2020A1515010680).

Acknowledgments: The authors would like to thank Xiaoping Liu from Sun Yat-sen University for providing the datasets. Furthermore, the authors would like to thank the editors and anonymous reviewers, whose detailed comments and suggestions have notably helped to improve the paper.

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

References

- 1. Farjam, R.; Motlaq, S.M.H.; Rasoul, F. Does urban mixed use development approach explain spatial analysis of inner city decay? *J. Urban Manag.* **2019**, *8*, 245–260. [CrossRef]
- 2. Li, M.; Koks, E.; Taubenböck, H.; van Vliet, J. Continental-scale mapping and analysis of 3D building structure. *Remote Sens. Environ.* **2020**, 245, 111859. [CrossRef]
- Atkins, K.G. Analysis of Town Center Mixed-Use Developments to Determine Key Retailer Success Factors. Master's Thesis, University of Tennessee, Knoxville, TN, USA, 2005.
- Bell, J. A Mixed-Use Renaissance: A Renewed Interest in Creating a Vibrant Urban Feel in City Centers and in Suburban Set-tings Is Propelling Successful Mixed-Use Development. Here Are Three Projects That Blend the Best of Mixed Use. *Infect. Immun.* 2004, 67, 5427–5433.

- 5. Dobbins, M. Urban Design and People; John Wiley & Sons: Hoboken, NJ, USA, 2011; p. 5.
- 6. Joe, G.A. A Vertical Vision; National Real Estate Investor: New York, NY, USA, 2004; pp. 24–30.
- Hoppenbrouwer, E.; Louw, E. Mixed-use development: Theory and practice in Amsterdam's Eastern Docklands. *Eur. Plan. Stud.* 2005, 13, 967–983. [CrossRef]
- Lee, J. Quality of Life and Semipublic Spaces in High-Rise Mixed-Use Housing Complexes in South Korea. J. Asian Arch. Build. Eng. 2011, 10, 149–156. [CrossRef]
- 9. Lu, Z.; Im, J.; Rhee, J.; Hodgson, M. Building type classification using spatial and landscape attributes derived from LiDAR remote sensing data. *Landsc. Urban Plan.* **2014**, *130*, 134–148. [CrossRef]
- 10. Huang, X.; Liu, H.; Zhang, L. Spatiotemporal Detection and Analysis of Urban Villages in Mega City Regions of China Using High-Resolution Remotely Sensed Imagery. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 3639–3657. [CrossRef]
- 11. Belgiu, M.; Tomljenovic, I.; Lampoltshammer, T.J.; Blaschke, T.; Höfle, B. Ontology-Based Classification of Building Types Detected from Airborne Laser Scanning Data. *Remote Sens.* **2014**, *6*, 1347–1366. [CrossRef]
- 12. Hecht, R.; Meinel, G.; Buchroithner, M. Automatic identification of building types based on topographic databases—A comparison of different data sources. *Int. J. Cartogr.* **2015**, *1*, 18–31. [CrossRef]
- 13. Huang, Y.; Zhuo, L.; Tao, H.; Shi, Q.; Liu, K. A Novel Building Type Classification Scheme Based on Integrated LiDAR and High-Resolution Images. *Remote Sens.* 2017, *9*, 679. [CrossRef]
- 14. Li, Y.; Huang, X.; Liu, H. Unsupervised Deep Feature Learning for Urban Village Detection from High-Resolution Remote Sensing Images. *Photogramm. Eng. Remote Sens.* **2017**, *83*, 567–579. [CrossRef]
- 15. Sritarapipat, T.; Takeuchi, W. Building classification in Yangon City, Myanmar using Stereo GeoEye images, Landsat image and night-time light data. *Remote Sens. Appl. Soc. Environ.* 2017, *6*, 46–51. [CrossRef]
- 16. Wurm, M.; Schmitt, A.; Taubenbock, H. Building Types' Classification Using Shape-Based Features and Linear Discriminant Functions. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 1901–1912. [CrossRef]
- 17. Chen, W.; Huang, H.; Dong, J.; Zhang, Y.; Tian, Y.; Yang, Z. Social functional mapping of urban green space using remote sensing and social sensing data. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 436–452. [CrossRef]
- 18. Chen, Y.; Liu, X.; Li, X.; Liu, X.; Yao, Y.; Hu, G.; Xu, X.; Pei, F. Delineating urban functional areas with building-level social media data: A dynamic time warping (DTW) distance based k -medoids method. *Landsc. Urban Plan.* **2017**, *160*, 48–60. [CrossRef]
- 19. Zhong, C.; Huang, X.; Arisona, S.M.; Schmitt, G.; Batty, M. Inferring building functions from a probabilistic model using public transportation data. *Comput. Environ. Urban Syst.* **2014**, *48*, 124–137. [CrossRef]
- Zhuo, L.; Shi, Q.; Zhang, C.; Li, Q.; Tao, H. Identifying Building Functions from the Spatiotemporal Population Density and the Interactions of People among Buildings. *ISPRS Int. J. Geo-Inf.* 2019, *8*, 247. [CrossRef]
- Zhao, Y.; Chen, G.; Chen, G.; Liu, X.; Niu, N. Integrating Multi-source Big Data to Extract Building of Urban Villages: A Case Study of Tianhe District, Guangzhou. *Geogr. Geo-Inf. Sci.* 2018, 34, 3–13.
- 22. Chen, W.; Zhou, Y.; Wu, Q.; Chen, G.; Huang, X.; Yu, B. Urban Building Type Mapping Using Geospatial Data: A Case Study of Beijing, China. *Remote Sens.* 2020, *12*, 2805. [CrossRef]
- 23. Zhang, Y.; Li, Q.; Tu, W.; Mai, K.; Yao, Y.; Chen, Y. Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. *Comput. Environ. Urban Syst.* 2019, 78, 101374. [CrossRef]
- 24. Liu, X.; He, J.; Yao, Y.; Zhang, J.; Liang, H.; Wang, H.; Hong, Y. Classifying urban land use by integrating remote sensing and social media data. *Int. J. Geogr. Inf. Sci.* 2017, 31, 1675–1696. [CrossRef]
- 25. Niu, N.; Liu, X.; Jin, H.; Ye, X.; Liu, Y.; Li, X.; Chen, Y.; Li, S. Integrating multi-source big data to infer building functions. *Int. J. Geogr. Inf. Sci.* 2017, 31, 1–20. [CrossRef]
- Liu, X.; Niu, N.; Liu, X.; Jin, H.; Ou, J.; Jiao, L.; Liu, Y. Characterizing mixed-use buildings based on multi-source big data. *Int. J. Geogr. Inf. Sci.* 2018, 32, 738–756. [CrossRef]
- 27. Kolda, T.G.; Bader, B.W. Tensor Decompositions and Applications. SIAM Rev. 2009, 51, 455–500. [CrossRef]
- 28. Sun, L.; Axhausen, K.W. Understanding urban mobility patterns with a probabilistic tensor factorization framework. *Transp. Res. Part B Methodol.* **2016**, *91*, 511–524. [CrossRef]
- 29. Tao, H.; Wang, K.; Zhuo, L.; Li, X. Re-examining urban region and inferring regional function based on spatial-temporal interaction. *Int. J. Digit. Earth* **2018**, *12*, 293–310. [CrossRef]
- Mørup, M. Sparse Higher Order Non-negative Matrix Factorization. Available online: http://www2.imm.dtu.dk/pubdb/pubs/ 4718-full.html (accessed on 26 July 2020).