MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

8—26

The automatic detection and recognition of objective parts in real-world building image

Jin Haomin *
Institute of Industrial Science
University of tokyo

Masao Sakauchi†
Institute of Industrial Science
University of tokyo

Abstract

Now, we are constructing a City-streets Building Database System. Particularly, we want to use this system to recognize real-world building images and provide the necessery information about the building objects. Because in building image there are many objects, and only the parts of building objects are necessary for recognition, we have developed a method by which all the parts of building objects are first detected using a statistical algorithm, and then the image features are extracted from the parts. Here, we suggest that the texture and color features are extracted respectively using "wavelet transform" and color-clustering. At the same time, a method called "vote" is used to decide the candidates for building object. Finally, we will give some experiment results to reveal the efficacy and accuracy of the proposed method in this paper.

1 Introduction

At present, based on the development of multimedia technology, the multimedia systems for processing real-world information are being developed widely in the world. The systems are needed not only to be able to effectively manage the data related to real-world objects, but also to present the ways of automatically recognizing real-world objects using computer vision technology. Considering these needs, we are constructing a City-street Building Database using the data realted to building objects and building images. Particularly, based on the vision indexes of the database constructed using building images, the database can be accessed using real-world building images.

For realizing the objectives mentioned as above, this paper will propose a method of automatically detecting and recognizing all the parts of building objects in building image. Here, the method is suggested to first detect all the parts of building objects using a statistical algorithm, and then the texture and color features are extracted respectively based on "wavelet transform" theory and color clustering method for recognition. In experiments, the real-world building images taken with a different angle, light, and scale on some extent from the image patterns of the database, are used to test the accuracy and efficacy of the method.

This paper is organized as follows. The detection of the building parts using the statistical algorithm is discussed in section 2. The extraction of color and texture features are given in section 3. The experiment results are presented in section 4. Finally, we summarize this paper and presented the future research work.

2 The detection of the building parts in real-world building image

For detecting the parts of building objects, we first detect the points representing texture information.

2.1 The detection of the points representing texture information

According to the practical needs, here the Y color component representing the luminance channel of CIE-XYZ color space[1] and a 3×3 window are employed to detect the points representing texture information. If the Y component value of the point located at the center of the window is the largest or the least within the window and the absolute difference value between the center point and other a certain point in the window is larger or smaller than a specilized threshold, the center point is detected as a point representing texture information(Fig.1).

2.2 Removing the noise points and recovering the points representing texture information

Although the points representing texture information can be detected correctly by the above-mentioned method, there are still many noises in the points representing texture information. This will make great influence to detecting the parts of

*Address: 7-22-1 Roppongi, Minato-ku, tokyo 106 Japan. E-mail: jhm@sak.iis.u-tokyo.ac.jp

†Address: 7-22-1 Roppongi, Minato-ku, tokyo 106 Japan. E-mail: sakauchi@sak.iis.u-tokyo.ac.jp

Figure 1: A building image and the points extracted using the method of detecting texture information

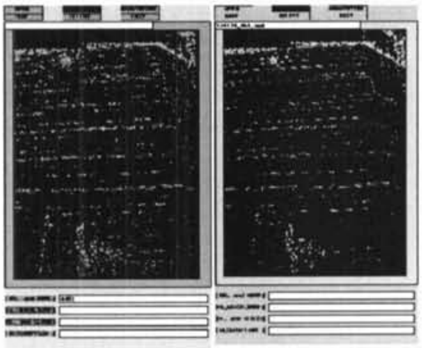

Figure 2: A original image of the points exracted from building image and the image of the points smoothed using the templetes.

building objects afterword. Therefore, it is necessary to remove these noise points and recover the points representing texture information. In this system, some 3×3 and 4×4 templates have been used to do these. A result example using this method is shown in Fig.2.

2.3 Deciding the ranges of building parts along horizontal direction

Here, the ranges of building parts along horizontal direction are detected using a statistical algorithm. The algorithm are shown as follows:

1. Compute the histogram of the points representing texture information distributing along the direction of horizontal axis. For finding easily the really variance positions, it is necessary to smooth the histogram. Here, at every position of the histogram, the average values of five-neighborhood is calculated to be the new values of the histogram as defined in equation (1).

$$H'(i) = \frac{\sum_{j=-5}^{5} H(i+j)}{11} \quad (1)$$

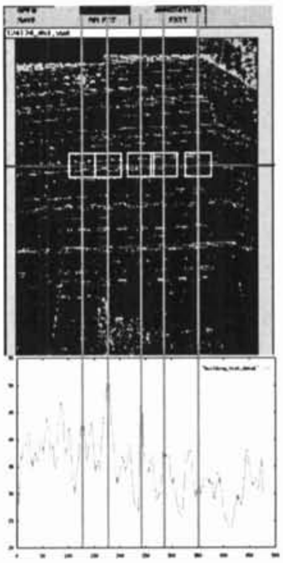

Figure 3: Clustering peaks based on the distribution of the points locating in the particular windows

$$P(i) = \begin{cases} 1 & H'(i) = maxH'(i+j), j = -5, \cdots, +5 \\ 0 & otherwise \end{cases} \quad (2)$$

2. Find all the positions of peaks (five-neighborhood maximum) using equation (2).

3. Cluster these peaks. Here, at every position of the peaks, a window with the size of 50×50 pixel is put up in the middle of image for clustering(Fig.3). The window is further divided into four small 25×25 sub-windows, and the proportions of the number of points locating in the 50×50 window to the number of the points locating in every the sub-window, are calculated respectively using equation(3). These four proportions are used to calculate all the distances of similarity between a certain peak and the right one from left to right by equation (4) for peaks clustering. If a distance is less than a specialized threshold, the corresponding two peak positions are considered belonging to the same building object.

$$P_n^i = [\frac{\sum_{u,v} T^i(u,v)}{\sum_{x,y} T_n^i(x,y)} + 0.5] \quad (3)$$

Here, $(u,v) \in$ window locating in peak i. $(x,y) \in$ the nth sub-window of the window locating at peak i. $T^i(u,v)$ equals one if the point locating at (u,v) is the one representing the texture information, otherwise zero.

$$D(i,j) = \sqrt{\sum_{n=1}^{4} (P_n^i - P_n^j)} \quad (4)$$

P_n^i : The proportion of the number of texture information points held by the window locating at the position of peak i to the one held by its nth sub-window. Here n=1,2,3,4.

4. If a certain peak and its right first-neighboring peak do not belong to the same building object, the distance between it and its right second-neighboring peak is calculated as step 3. If the distance is less than a specialized threshold, and the distance between it and its first-neighboring one is less than a specialized threshold too, these three peak positions are regarded belonging to the same building object.
5. Decide the ranges of building parts. For doing this, here if the number of the peaks belonging to a certain cluster is one, the specialized neighboring range centered round the peak position is considered to be the range. Otherwise the range between the first peak and the last peak belonging to the cluster is considered as the range.
6. Finally considering the situation that the original building part is still possibly treated to be some different building parts, the ranges decided using the above steps are further merged. The method is that if the any two ranges overlap with each other, these two ranges will be merged into one and some new ranges are defined.

Based on the above steps, all the ranges of building parts along horizontal direction can be decided.

2.4 Deciding the ranges of building parts along vertical direction

Deciding the ranges of building parts along vertical direction is done respectively within every the range decided by the method mentioned in 2.3. Concretely, within every the range along horizontal direction, the histogram of the points distributing along the direction of vertical axis are computed. The peak positions of the histogam are detected using the same method described in 2.3. The peaks which value are more than a specialized threshold are detected, and the two peaks locating at most top and bottom positions are selected to decide the range of building parts along vertical direction.

3 The extraction of image features

Considering the needs of constructing City-streets Building Database, the texture features are extracted using wavelet transform [3]. Wavelet transform has been successfully applied to many image analysis tasks such as edge detection, corner detection, texture classification, object recognition, image segmentation, and shape recovery. In our system, we have applied this technique to resolve the problem of building image recognition. Color is also very important recognition feature, which have been researching widely in many application fields. Because the color variance of building image occurs generally among the glass, wall and window frame, the points of texture information can been detected from the variance of image color. Therefore, if the color features are extracted from these points representing texture information, it will be efficient and effective. Here, first the colors of the points representing texture information detected using the method proposed in 2.1, are transformed from the RGB color space to the YES color space[2]. Secondly we employ the E,S color components to compute the color histogram of the points representing texture informationin. Finally, Based on the ES color histogram, the colors are clustered in ES color space. The E, S color components of the clusters obtained by clustering colors are used as the color features for recognition.

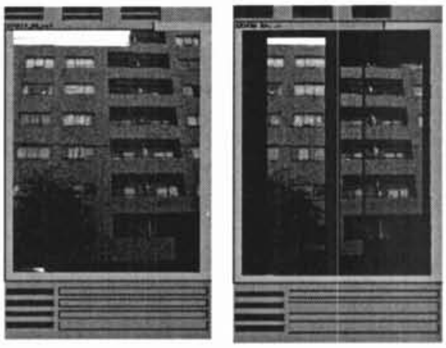

Figure 4: An example of deciding the ranges of the objective buildings

4 Experiments

4.1 Experiment for detecting the parts of building objects

In experiments, 200 building images are used to test the method of detecting the parts of building objects proposed here. By the experiments it can be observed that the different building objects and the building parts with different structure can be detected in real-world building image. The Fig.4 is an example of automatically detecting the parts of building objects using the method proposed in this paper.

4.2 The recognition based on "vote" and experiment

We first constructed City-streets Building Database of 200 building images. For constructing the indices of the building images, the part of building image including less noise and being able to represent the vision

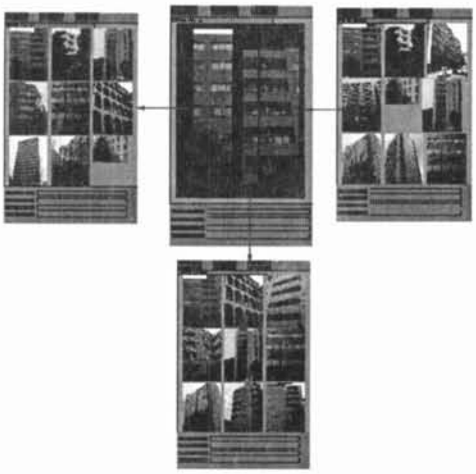

Figure 5: Retrieving the database using the features of sub-window image

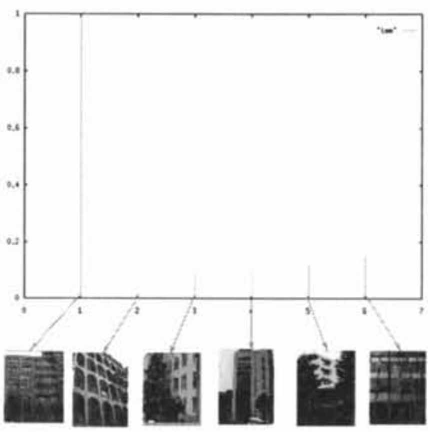

Figure 6: Deciding the candidate buildings based on the vote numbers

characters of building objects are first marked manually, and then the features are extracted from the part using the method metioned in 3. The recognition method based on the "vote" is used to recognize the other building images different from the building image patterns of the database. The algorithm using "vote" to recognize the building image is as follows:

1. Retrieve the database respectively using the features extracted from every the part of building object(Fig.5), and obtain several lists of limited candidates for the parts of building object according to the distances of similarity.
2. Compute the vote values of the candidates according to the positions in every the list of limited candidates.
3. Decide the candidates for the building object based on the vote values.

Table 1: The recognition rate(%)by marking real-world building images

Distance list	Recognition rate
Position I	65.7
Position II	82.8
Position III	100

Position I : The top one position.
Position II : Within the top five positions.
Position III: Within the top ten positions.

When retrieving the database using the building images, first the parts of building objects are detected, and then the features of texture and color are extracted from these parts. For every the part, the texture features are first used to retrieve the database and a list of limited candidates for the building object is obtained, then the color features are used for further deciding the positions of the limited candidates in the list. Here the experiment is done using the building images that are different from the image patterns of the database in angle, light, and scale on some extents. The recognition rate is shown in Table 1. From the experiment result it can be observed that the objective building image can be retrieved within the limit candidates, although the vision factors have some changes to some extent.

5 Summarize

This paper proposed the method of automatically detecting and recognizing the parts of building objects in real-world building image. The efficacy of the method is revealed by experiments. But the method on the detection have too its inadequate sides. For example, the method will not be able to detect the part of building object where the texture features do not exist. In future, the other image features will be researched for the detection and recognition of building objects. At the same time, we are also going to research the method of recognizing building objects based on the adjacent relations among building objects.

References

[1] Yihong Gong, Masao Sakauchi, "Detection of Region Matching Specified Chromztic Features", Computer vision and image understanding, Vol. 61, No. 2, Match, pp. 263-269, 1995.

[2] Eli Saber, A. Murat Tekalp, Reiner Eschbach, Keith Knox, "Automatic Image Annotation Using Adaptive Color Classification", Graphical models and image processing, Vol. 58, No. 2, Match, pp.115-126, 1996.

[3] Stephane Mallat and Sifen Zhong, "Characterization of Signal from multiscale Edges", IEEE Tansaction on PAMI, Vol. 14, No. 7, 1992.