Surveillance and Recognition of Media Information

Hiroshi Murase, Yoshito Mekada,
Takashi Kusanagi, Osanori Matsugano

Department of Media Science, Graduate School of Information Science

murase@is.nagoya-u.ac.jp

We are currently developing media surveillance and recognition systems. Surveillance technologies that monitor long-term video sequences or vast amounts of image data to obtain useful information are important for realizing the safety and comfort of humans. It is, however, not easy to quickly and accurately analyze video/image data captured in real environments. We have approached this problem from the angles of both speed and accuracy, and here we introduce two results from our recent researches.

Object Surveillance Using Multiple Cameras

Imagine a surveillance system in a common situation. The system should be able to recognize a small object (e.g., a human face) by using distant cameras under unspecified illumination conditions. The orientation of the object is not fixed. In this situation, the quality of the image is not sufficient to guarantee accurate recognition. However, even if we cannot recognize the object from such a low quality single image, it might be possible to recognize the object when using multiple images captured from different viewpoints. We have developed a new object recognition method based on Parametric Eigenspace Representation for such multiple viewpoint images. In this method, for each object of interest, a large set of images obtained from various positions is compressed using the principal component analysis to get a low-dimensional subspace called the eigenspace. The object is represented as a hyper-surface in this space. Then, given a set of unknown input images from multiple viewpoints, the recognition system projects the images to the eigenspace. The object is recognized based on the distances between the projected points and the hyper-surface. We use the relative positions of the cameras as a restriction condition when measuring such distances in the eigenspace. We demonstrate that the recognition rate is improved by using multiple cameras compared with a single camera. We also investigate the relation between the arrangement of the cameras and the classification performance.

Medical Surveillance for CT Images

It is important for human health care to surveil one's condition over the long term. We have developed a method to find corresponding nodules between previous and current CT images. To measure the curative effect of lung nodules quantitatively, it is necessary to find the corresponding nodules between images. These corresponding pairs and their accompanying statistics may be good indices for determining whether current medical treatment is effective. To achieve this, we have developed a computer-aided diagnosis algorithm to detect hundreds of small nodules from CT images and find corresponding pairs in order to evaluate the curative effect for comparative reading. We employ relative distance for coarse matching and image correlation between volumes of interest for precise matching. Our method consists of the following parts: extraction of lung area for reducing both false positives and computational time; detection of small nodules by using two types of shape features; coarse matching by using the relative position of each nodule in the lung area; and precise matching using local gray value features. We applied the method to four chest CT images that were taken during successive examinations of the same patient in one year. It could match correspondences to an accuracy of 95%. Because several features could be measured easily, such as the total number and volume of nodules, appearance of nodules, disappearance of nodules, it is expected that this method will be an effective diagnostic tool.

Surveillance and Recognition of Media Information

Hiroshi Murase, Yoshito Mekada, Takashi Kusanagi, Osanori Matsugano
Department of Media Science, Graduate School of Information Science

Surveillance technologies

Mining useful information quickly and accurately from video/image data captured in real environments.

e.g. { Long-term video sequences
Vast amounts of image data

Objectives

To develop fundamental pattern recognition algorithm for media information and their applications.

To realize the safety and comfort for people.

Keywords

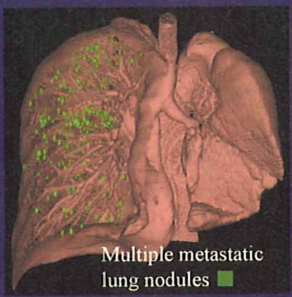
Multiple-viewpoint images, Parametric Eigenspace method, Long-term video sequences, Low-resolution images, Medical image analysis

Medical Surveillance for CT Images

Background

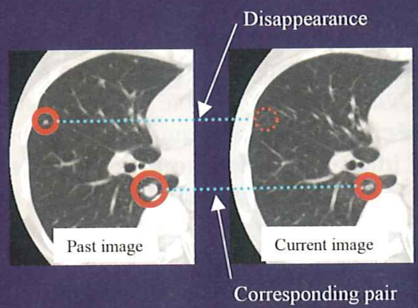
Follow-up surveillance of multiple lung nodules over the long term.

- Hundreds of slice images for comparative readings.
  - ➡ Heavy load on doctors.
- Many nodules.
  - ➡ Difficult to find the corresponding nodule.



Purpose

- To find corresponding individual nodules
- To evaluate temporal changes in lung nodule properties.
  e.g.: the number of nodules, appearances, disappearances



Coarse-to-fine matching method

Coarse matching (using relative position)

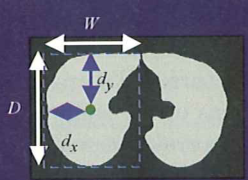
1) The relative position of each nodule in the lung area is calculated.
2) The distance between two nodules is calculated for all combinations of nodule pairs on the same side of a lung.

Precise matching (using image correlation)

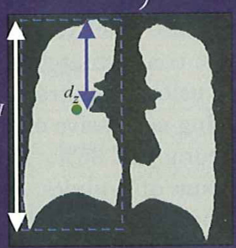
1) Neighborhood area of each nodule is projected by Maximum Intensity Projection (MIP) method.
2) Image correlations are calculated for all nodule pairs that are close to each other.

The relative position of nodule A_i

$$P_{A_i}(x_i, y_i, z_i) = \left(d_x/W,\ d_y/D,\ d_z/H \right)$$



Axial section

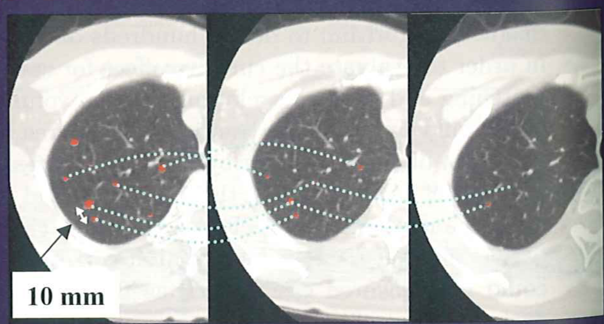


Coronal section

The distance between nodule A_i and nodule B_j

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

Experimental results



Data 1 Data 2 Data 3

Conclusions

Detect 260 pairs out of 265 with only 12 false positives.

➡ Toward the quantitative evaluation of the curative effect.

Object Surveillance using Multiple Cameras

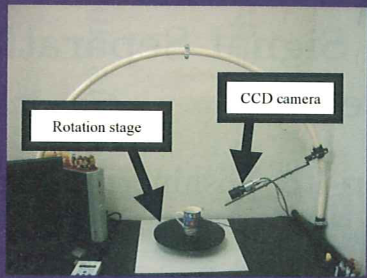
Background

- The quality of images captured by surveillance systems might be too low for recognition.
- Using multiple cameras improves the accuracy.

Purpose

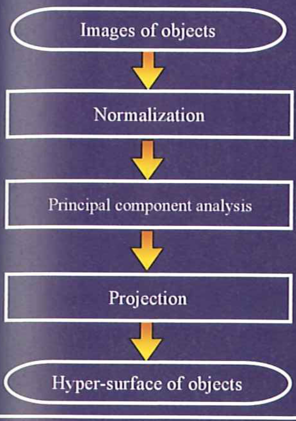
- The aim is to determine how many cameras are the most suitable and where should they be located.
- Parametric Eigenspace Representation is used.

An image acquisition device for learning



It is able to capture an object from arbitrary view angles automatically.

Learning process

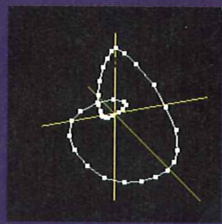


A large set of images obtained from various positions is compressed using the principal component analysis to the eigenspace. Each object is represented as a hyper-surface in the eigenspace.

Learning samples $\{x_i^c\}$
c: category number
i : sample number

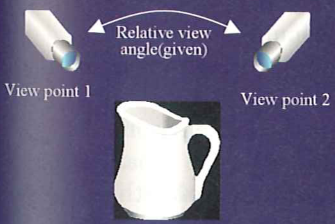
The principal component analysis of XX^T ($X=\{x_1^1, x_2^1, \ldots, x_n^c\}$)

Projection to the eigenspace



Hyper-surface in the eigenspace

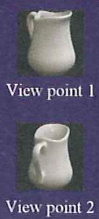
Recognition process



An object to be recognized

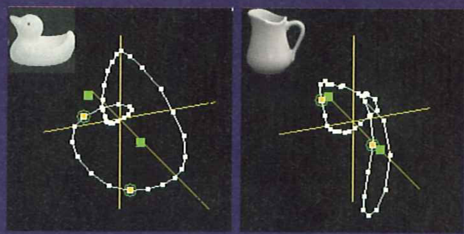
The object is recognized based on the distances between the projected points and the hyper-surface.
The relative positions of the cameras is used as a restriction condition when measuring distances in the eigenspace.

Image acquisition



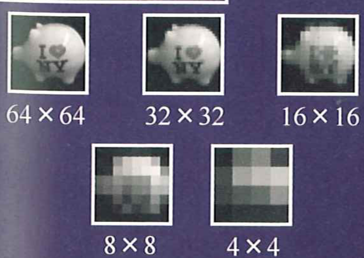
Input images

Projection to the eigenspace

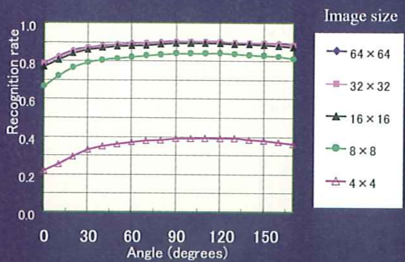


Finding the minimum distance between hyper-surface and the projected points (■)

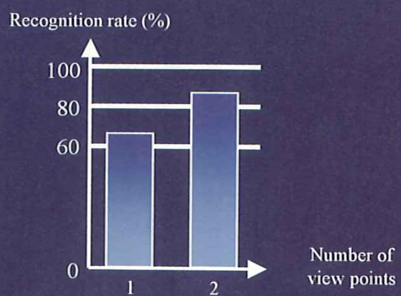
Experimental results



Examples of input images. Image size is described below



Recognition rate vs. Relative viewing angle.



Average recognition rate vs. Number of view points

Conclusions

The recognition rate is improved by using multiple cameras instead of a single camera.
The most desirable relative viewing angle between two cameras is around 115°.