ISSN 0002-306X. Изв. НАН РА и ГИУА. Сер. ТН. 2010. Т. LXIII, №3.

UDC 004.032.26:629.3.066.8.52 AUTOMATION AND CONTROL SYSTEM

G.A. POGHOSYAN

DECREASING VOLUME OF IMAGE DATABASE FOR AUTOMATION FACE RECOGNITION SYSTEM

Granting computers ability to discover and identify the object is a contemporary problem that deserves special attention. This problem is widely available in various technological processes, security services, reading systems, criminal law, medical systems, video-conferences, in computer virtual reality, as well as in all video observation systems. Because the recognition database is a collection of images and automatic face recognition system should work with these images, which can hold large volumes of computer memory that is way it's necessary to investigate and develop a method / tool for optimal using volume of computer memory (that decrease image a database volume) and implement quick search within database.

Keywords: Haar-like features, LEM (Line Edge Map), invariant, pixel, frame.

Investigations and study of certain methods help us to develop and implement the method/tool, which decreases volume of recognition database. To achieve the above mentioned goal the image LEM (Line Edge Map) is used [1]. Since quantity of images in recognition database (n) is not infinite, images (*.bmp, *.jpg), n=40 and *.jpg format are selected. For object detection and recognition algorithm Haar-like features are used. They owe their name to their intuitive similarity with Haar wavelets and are used in the first real-time face detector.

Image sensors have become more significant in the information age with the advent of commodity multi-media capture devices such as digital cameras, webcams and camera phones. The data from these media sources (whether they are still images or video) reach the stage where manual processing and archiving become impossible. It is now possible to process these images and videos for some applications in near real-time using motion detection and face tracking for security systems. However there are still many challenges including the ability to recognize and track objects at arbitrary rotations. Haar-like features have been successfully used in image sensors for face tracking and classification problems [2], however other problems such as hand tracking [3] have not been so successful. Historically, working with only image intensities (i.e., the RGB pixel values and every pixel of image) made the task of feature calculation computationally expensive. A publication by Papageorgiou [4] discussed an alternate feature set based on Haar wavelets instead of the usual image intensities. Viola and Jones [5] adapted the idea of using Haar wavelets and developed the so-called Haar-like features. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in these regions and calculates the difference between them. This difference is then used to categorize subsections of an image. For example, we have an image database with human faces. It is a common observation that among all faces the region of the eyes is darker than the region of the cheeks. Therefore a common Haar feature for face detection is a set of two adjacent rectangles that lie above the eye and the cheek region. The position of these rectangles is defined relative to a detection window that acts like a bounding box to the target object (the face in this case).

In the detection phase of the Viola-Jones object detection framework, a window of the target size is moved over the input image, and for each subsection of the image the Haar-like feature is calculated. This difference is then compared to a learned threshold that separates non-objects from objects. Because such a Haar-like feature is only a weak learner or classifier (its detection quality is slightly better than random guessing) a large number of Haar-like features are necessary to describe an object with sufficient accuracy. In the Viola-Jones object detection framework, the Haar-like features are therefore organized in something called a classifier cascade to form a strong learner or classifier.

The key advantage of a Haar-like feature over most other features is its calculation speed. Due to the use of integral images, a Haar-like feature of any size can be calculated in constant time (approximately 60 microprocessor instructions for a 2-rectangular feature).

The main reason for this is the fact that Haar-like features are not invariant over rotation. This means that any object that rotates and is sensitive to angle changes (such as hands) will be difficult to solve using standard Haar-like features. The features that define faces tend to be insensitive to small angle variations and Haar-like features have been used to detect head rotations of as much as 15^0 from the vertical [6]. When people are standing their heads is naturally aligned vertically with respect to gravity, this rotational sensitivity tends not to be a significant problem for faces. Other body parts such as hands, arms and legs are not normally aligned with the horizontal or vertical axes and they are difficult to model with traditional Haar-like features. Researchers have tended to use edge detection or color based tracking of these parts.

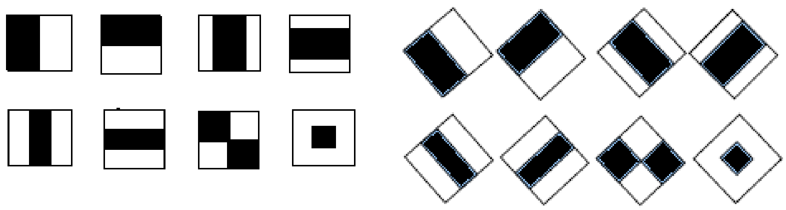

Fig. 1. Standard Haar-like features and 45^0 twisted Haar-like features

Several researchers have studied the impact of in plane rotations for image sensors with the use of twisted Haar-like feature (45^0) [6] or diagonal features of fairly good performance have been achieved. These techniques will have little benefit for problems that are sensitive to rotations, such as hand identification [3], which are not aligned to fixed angles (0^0, 45^0, 90^0, etc).

Haar-like feature based classifiers like the Jones and Viola, face detector work in almost real time using the integral image(or summed area table) data structure that allows features to be calculated at any scale with only 8 operations. However standard Haar-like features are strongly aligned to the vertical/ horizontal or 45^0 (Fig. 1.) and so are most suited to classifying objects that are strongly aligned as well, such as faces, buildings, etc.

Standard Haar-like features consist of a class of local features that are calculated by subtracting the sum of a subregion of the feature from the sum of the remaining region of the feature.

Integral images or summed area tables are a data structure that contain the sum of all the pixels above and to the left of the current pixel. The time complexity of the algorithm is 2MN (where M and N are the height and width of the image) since each pixel in the integral image requires two addition operations (see eq. 1).

$$f(i,j) = f(i-1,j) + f(i,j-1) - f(i-1,j-1) + g(i,j), \qquad (1)$$

where $g(i,j)$ is the pixel value at position (i,j), $f(i,j)$ is the integral image value at position (i,j). Integral images are important as they allow the sum of a rectangular area of pixels of any size to be calculated with only 4 look ups in the Integral image data structure:

$$\sum_{i=a}^{b}\sum_{j=c}^{d} g(i,j) = f(a,c) - f(a,d) - f(b,c) + f(b,d), \qquad (2)$$

where $g(i,j)$ is the pixel value at position (i,j), $f(i,j)$ is the integral image value at position (i,j), (a,c) is the coordinate of the top left pixel and (b,d) is the coordinate of the bottom right pixel of the rectangular region that is being summed.

Edge information is a useful object representation feature that is insensitive to illumination changes to certain extent. Edge detection is a fundamental tool in image processing and computer vision, particularly in the areas of feature detection and feature extraction, which aim at identifying points in a digital image at which the image brightness changes sharply. The purpose of detecting sharp changes in image brightness is to capture important events and changes in properties of the world. In the ideal case, the result of applying an edge detector to an image may lead to a set of connected curves that indicate the boundaries of objects, the boundaries of surface markings as well as curves that correspond to discontinuities in surface orientation. Thus, applying an edge detection algorithm to an image may significantly reduce the amount of data to be processed and may therefore filter out information that may be regarded as less relevant, while preserving the important structural properties of an image. If the edge detection step is successful, the subsequent task of interpreting the information contents in the original image may therefore be substantially simplified. However, it is not always possible to obtain such ideal edges from real life images of moderate complexity. Edges extracted from non-trivial images are often hampered by fragmentation, meaning that the edge curves are not connected, missing edge segments as well as false edges not corresponding to interesting phenomena in the image thus complicating the subsequent task of interpreting the image data [7].

Rectangular mask ($n \times n$) for getting LEM of image (Fig. 2) is used. All images having $n \times n$ size of rectangles (n=8) and these rectangles must be smaller than the main features of the image details (nose, mouth, eyes). It helps us to observe all brightness points and avoids not important pixels of image.

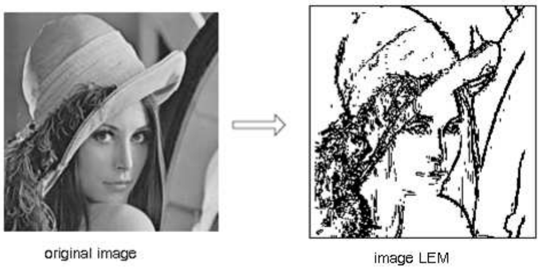

Fig. 2. Used 8x8 mask for getting image LEM

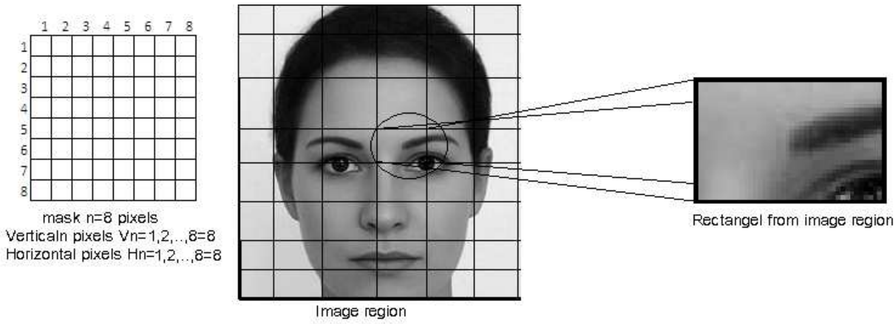

Fig. 3. 8x8 mask and image divided with 8x8 rectangles

Equation (3) is used for calculating brightness of horizontal or vertical pixels and their opposite side pixels (Fig. 3).

$$V = \sum_{i=1}^{n} (f(i,1) - f(i,8) , \qquad H = \sum_{j=1}^{n} f(1,j) - f(8,j) . \tag{3}$$

The function having brightness values of horizontal and vertical sides on current rectangle, which shows lines according to brightness values (Fig.3) is developed.

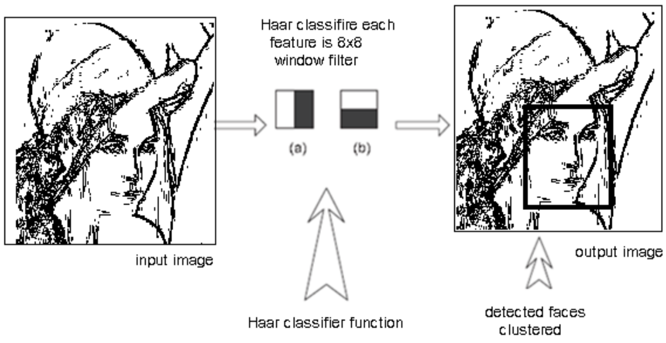

Fig. 4. Haar-classifier face detection algorithm data flow

Fig. 4 shows changing recognition images data flow for Haar classifier face detection algorithm. Using black and white liner image method for training database the following results are given in Table, where t_{RGB} and t_{BW} are face detection and recognition time, S_{RGB} and S_{BW} are sizes of training images are used for recognition.

Table

Training database type	Training images	Detection and recognition time	Training database size
RGB	40	t_{RGB}	S_{RGB}
BW(black and white)	40	t_{BW}	S_{BW}

After many experiments we get the following results:

$$t_{RGB} = t_{BW} \quad \text{and} \quad S_{BW} = S_{RGB} \times (0.6 - 0.58), \tag{4}$$

which means that image database size is reduced for 40÷42 % percentage. So we get image optimization tools which can help us to optimize and reduce the file size of the image database for recognition and developed a method / tool for optimal using volume of computer memory .

So in future work we will try to implement real time frame moving face discovering and identifying the automation system, which will use the result of this work and will try to identify and recognize faces using database with LEM of images.

REFERENCES

1. **Tolba A. S., El-Baz A.H., and El-Harby A.A.** Face Recognition //A Literature Review International Journal of Signal Processing. – 2006.-P. 2.
2. **Koutroumbas Konstantinos, Theodoridis Sergios.** Pattern Recognition (4th ed.). Boston: Academic Press, 2008, ISBN 1-59749. -P. 272-8.

3. **Shan C., Wei Y., Tan T., Ojardias F.** Real time hand tracking by combining particle filtering and mean shift // Proc. Sixth IEEE Automatic Face and Gesture Recognition, FG04.- 2004. -P. 229-674.
4. **Papageorgiou B.** A general framework for object detection // International Conference on Computer Vision. -1998.
5. **Viola P. and Jones M.** Rapid object detection using a boosted cascade of simple features // Presented at 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. -2001.-Vol.1.-P. 511-518.
6. **Mita T., Kaneko T., and Hori O.** Joint Haar-like features for face detection // Presented at Computer Vision Tenth IEEE International Conference on. -2005. -Vol. 2. –P. 1619-1626.
7. **Lindeberg T**. Edge detection and ridge detection with automatic scale selection // International Journal of Computer Vision. -1998. -P.117-154.

Institute for Informatics and Automation Problems of NAS of RA. The material is received on 12.09.2010.

Գ.Ա. ՊՈՂՈՍՅԱՆ

ԴԵՄՔԻ ԱՎՏՈՄԱՏ ՃԱՆԱՉՄԱՆ ՀԱՄԱԿԱՐԳԻ ՊԱՏԿԵՐՆԵՐԻ ՏՎՅԱԼՆԵՐԻ ԲԱԶԱՅԻ ԾԱՎԱԼԻ ՆՎԱԶԵՑՈՒՄ

Համակարգչին օբյեկտները նկատելու հատկությամբ և ճանաչողական ունակությամբ օժտելը այսօր խիստ արդիական խնդիր է և արժանի է ուշադրության: Այն առկա է տեխնոլոգիական գործընթացներում, պահակային ծառայություններում, քրեագիտության մեջ, և տեսահսկում կատարող բոլոր այլ համակարգերում: Այս աշխատանքում կատարված են այդ խնդիրների լուծման հայտնի մեթոդների ուսումնասիրություններ և հետազոտություններ: Մեր նպատակն է հետագա մեր աշխատանքներում մշակել և իրականացնել դեմքի ճանաչման և վերլուծության համակարգ, որը կօգտագործի այս աշխատանքում ներկայացված գծային պատկերների բազան :

Առանցքային բառեր. Հաարի բնութագրիչներ, գծային պատկեր, ինվարիանտ, պիկսել, կադր:

Г.А. ПОГОСЯН

УМЕНЬШЕНИЕ ОБЪЕМА БАЗЫ ДАННЫХ ИЗОБРАЖЕНИЙ В СИСТЕМЕ АВТОМАТИЗАЦИИ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦА

Способность компьютеров обнаруживать и идентифицировать объект - актуальная проблема, которая заслуживает внимания, особенно в различных технологических процессах, охранных услугах, системах считывания, медицинских системах, в ходе видеоконференции, в компьютерной виртуальной реальности, а также во всех видеосистемах наблюдения. Поскольку база данных - это коллекции изображений и автоматическая система распознавания лица должна работать с этими изображениями, занимающими большой объем памяти компьютера, то необходимо исследовать и разработать метод - инструмент для оптимального использования объема памяти компьютера (уменьшение объема базы данных изображений) и осуществить быстрый поиск в базе данных.

Ключевые слова: определители Гаари, линейное отображение, инвариантность, пиксел, кадр.