

COLOUR, TEXTURE AND SHAPE FEATURE ANALYSIS FOR PERSON RE-IDENTIFICATION TECHNIQUE

Poongothai¹ and Suruliandi²

^{1,2}Department of Computer Engineering,
Manonmaniam Sundaranar University, Tamilnadu, India

ABSTRACT

Person re-identification across different surveillance cameras has become one of the recent and challenging approaches in the area of video surveillance system. Although many methods have been introduced for re-identification, certain issues like illumination, pose variation and occlusion are unsolved. All the existing methods focused only on developing learning methods by varying the features. The aim of this work is to analyze the effectiveness of feature for re-identification. The colour, texture and shape features has taken for the analysis. The images are collected from multiple databases for training and testing. The analysis shows that every feature has its unique properties and works according to that to solve the issues

KEYWORDS

Re-identification, Colour, Texture, Shape ,Surveillance

1. INTRODUCTION

A fundamental task of a distributed multi-camera surveillance system is to identify the people across camera views at different locations and times. This is known as the person re-identification (re-id) problem and it is used in many crucial applications such as Criminal identification, video surveillance, access control, ID Card and forensic search . The purpose of re-identification system is to accurately recognize a person's identity from the cameras in a distributed network and take some actions on the output of the recognition process. The classical biometric system is solving this problem using finger print, palm, iris and face [8]. But these features cannot be use for automatic environment like surveillance network in public places. It is necessary to concentrate on the full appearance of persons.

Achieving automated re-identification remains a significant challenge due to the limitations of below issues [7]. Pose and viewpoint variations is the main challenge of a person with respect to the cameras of the network. It varies depending on the walking path of that person, and of the viewpoint of the camera. This may cause consistent variations of the person appearance. Another issue is occlusion. The parts of a person may be not visible to the camera due to occlusions caused by objects, clothing accessories or other people. This may cause the segmentation algorithm to fail in separating one person from the rest of the scene; consequently, descriptors may be built from images partially corrupted by the source of the occlusion. Illumination conditions may varied based on camera, in the same camera in different periods of time due to changing environmental conditions. This may result in appearance changes over different cameras and during time. Changes in colour response create the significant change in accuracy. Different cameras may have a different colour response that may affects the person appearance as well.

The above said issue hardly affects the re-identification accuracy. The re-identification system is the combination of multiple subroutines such as body part detection, feature extraction from parts, similarity matching and classification. The overall re-identification accuracy can be improved by introducing an efficient algorithm for the subroutines. This paper focuses on the main part of re-identification such as feature analysis. To understand the properties of all type of feature this paper concentrate on three type of common feature such as colour texture and shape. For colour RGB, for texture SIFT and for shape canny edge detection. From this we can analyse the performance of features for re-identification.

1.1. RELATED WORK

In recent years the research intense in person re-identification is keep on increasing by improving various learning methods and feature set. Some of the formidable researches and the ideas of person re-identification in recent years are listed below.

Bedagkar *et al* [2] made a brief survey and explore the person re-identification problem. In that research, available Re-ID methods, datasets and their pros and cons are elaborated. Unaddressed issues like open set Re-ID and long period Re-ID are also explored. Zhao et al [4] proposed a novel perspective for person re-identification based on unsupervised salience learning. Here the distinctive features are extracted using SIFT, RGB features without requiring identity labels in the training procedure. Bingpeng et al [5] proposed a new descriptor for person re-identification by building on the recent advances of Fisher Vectors. Specifically as pixel coordinates, its intensity as well as the first and second-order derivatives is computed for each pixel of the image. These local descriptors are turned from SIFT into Fisher Vectors to produce a global representation of the image. Martinel et al [6] given the detail description about some of the local features which are used for person re-identification in wide area network. They introduce the appearance based method for re-identification by extracting the local features from the appearance of the person image or video. It has proven to be robust against low resolution images, viewpoint and illumination changes, occlusions and pose variations. Cheng et al [9] presented the comparative study of local features for the task of person (re) identification based on the interest point detectors and descriptors GLOH and SIFT texture feature for re-identification. Doretto et al [10] made a review of all appearance-based person re-identification problems and approaches. This work focused on re-identifying people based on their whole body appearance.

1.2. MOTIVATION AND JUSTIFICATION

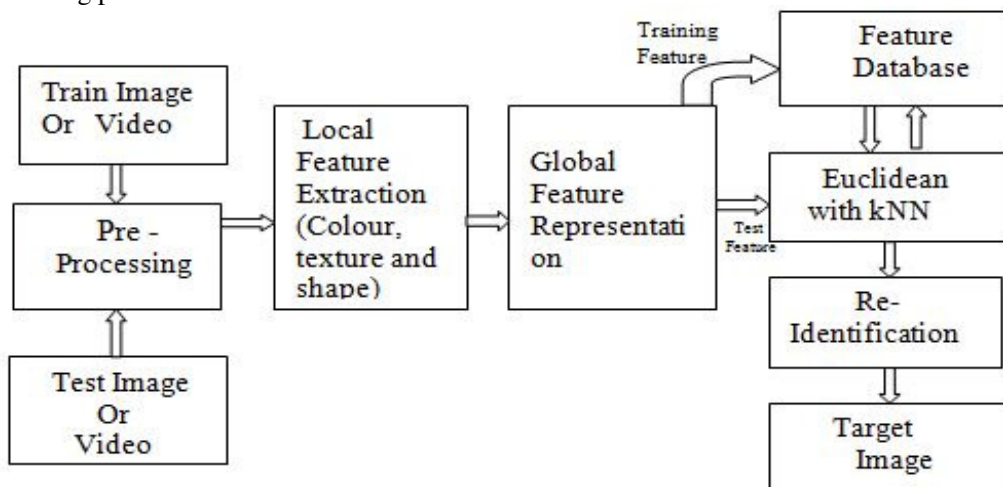
Surveillance system plays a significant role in security by identifying the unconventional actions of the people in public places. Now a day the events such as terrorist attacks have occurred often in many public places. The safety of people is become an emerging one in public places. So it needs to guarantee the safety of people in public places by detecting and preventing the violent actions. Person re-identification become the part of the process that is needed for the analyzing the activities and detect the person in surveillance network. Motivated by this concepts my research is focus to develop a new efficient algorithm for person re-identification to overcome all the issues.

To improve the efficiency of features for person re-identification it needs to analyze the performance of feature initially. There are many colour, texture and shape features available for re-identification. Among all the local features, it is observed from the literature that SIFT, RGB and shape context are the important basic level features for re-identification. Justified by this fact a comparative analysis on SIFT, RGB and shape context features is performed for various issues to validate the performance of colour, texture and shape feature. The analysis gives the detail about the importance of feature and can develop new feature from this for re-identification.

1.3. OUTLINE OF THE WORK

In this paper color, texture and shape features are experimented for person re-identification problem. The analysis of feature extraction is carried out in two phases namely training phase and testing phase. In both case pre-processing technique is used to resize images. During the training phase features are extracted from the training images and stored in training dataset. When the test image is given, the feature is extracted from the test image. Once a suitable feature representation has been obtained matching algorithm is used to measure the similarity between training images and test image. The frame work for feature extraction and the classification process is shown in Fig.1.

Training phase



Test phase

Fig.1. Block Diagram of Re-identification System

1.4. ORGANISATION OF THE PAPER

The remaining part of the paper is organized as follows: The feature extraction methods for person re-identification discussed in section 2. The procedure for re-identification in section 3. The results and analysis of all features are reported in section 4. The conclusion and future scope are summarized in section 5.

2. FEATURE EXTRACTION

2.1. PRE PROCESSING

Every method needs a pre-processing for further process. In this paper the input images which are taken from the database are in different size. So resizing and cropping is applied to all the input person images before extracting the features to convert into the common size and track the parts from person to extract the features.

2.2. FEATURE REPRESENTATION

Feature Extraction is the process of extracting the suitable features from the images as a distinguishable extent. Information extracted from images such as colour, texture and shape are known as feature vectors [8]. The extraction process is done on both query images and images in the database. Image matching involves based on the colour texture and shape features of both images and comparing them to search for similar features of the images in the database .

2.2.1 RGB COLOUR FEATURE

RGB colour information is used define the intensity of the person image based on RGB colour space where R, G and B are the values of the Red, Green and Blue colour channels. The colour feature is mainly extracted from the brightness ratio of three color channels. The colour ratio is computed by the eqn (1),(2),(3)

$$r = \frac{R}{R + G + B} \quad (1)$$

$$g = \frac{G}{R + G + B} \quad (2)$$

$$b = \frac{B}{R + G + B} \quad (3)$$

2.2.2 SCALE-INVARIANT FEATURE TRANSFORM (SIFT) TEXTURE FEATURE

SIFT, is a key point detector and feature descriptor which is appeared in all objects. The SIFT features are invariant to image scale and rotation variation and also robust to changes in illumination, noise, and minor affine transformations. It is computed as a histogram of the gradient distribution in the region around a detected interest point [9]. The gradient's orientation is selected based on the quantized regions which are centered on the key point resulting as the descriptors computed from magnitude. The magnitudes depend on the distance from the key point then, gradients are weighted by a Gaussian function. The descriptor is normalized in order to obtain illumination invariance.

2.2.3 CANNY EDGE BASED SHAPE FEATURE

Shape Context is mainly designed for detect the shape of the person image and matching is computed based on shape feature descriptor. It is an edge-based descriptor. Here the edges are detected based on the canny The descriptor consists of a histogram over the edge points, taking into account the location in 9 bins and edge orientation in 4 bins. The resulting descriptor is an extended version of the original Shape Context descriptor [7] and has 36 dimensions. $f(x,y)$ is the pixel value of the image which has occurred inside the edge of shape. We can view this as binary function in the domain space D.

$$f(x,y) = \begin{cases} 1 & \text{if } f(x,y) \in D \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

3. RE-IDENTIFICATION PROCEDURE

Formally, person re-identification can be modeled as a recognition/matching task [8], where a probe individual is matched against a gallery of templates (representing the individuals previously seen by the camera network). Thus, the problem of re-identifying an individual represented by its descriptor can be followed as:

3.1 TRAINING PHASE

1. The person is detected and tracked by using data association algorithm that track the people found by the detector to recognise the image regions, or blobs, that contain a person.
2. The pixels belonging to the person are separated from the background (foreground extraction or segmentation) in each frame of the video-sequence using an adaptive model of the background
3. The body is divided into sub parts; this phase can be carried out using a fixed part sub divisions.
4. Feature descriptor is built from the each body parts in the form of silhouettes (one for each frame) by using the global feature such as SIFT, RGB and canny edge based Shape context feature properties.
5. The feature of all local parts of person represented as feature vector.
6. Then the feature is represented globally for all the person into the account of single non rigid nature of the body.
7. The feature vector is generated from all the given images and stored in training data base

3.2. TEST PHASE

The same feature extraction procedure is followed for test image as it is in training phase. The feature has been represented for every given test image from the resulting silhouettes using global feature such as RGB, SIFT and Shape context. Then it is compared with the feature descriptors which are already trained and stored in training database.

3.3. MATCHING PROCEDURE

The similarity measure of a person in the training image with another person in the queried image is computed as the product of feature similarities. The feature matching of persons using RGB, SIFT and canny edge based shape context features is done through the Euclidean-distance based nearest neighbor approach. To increase robustness, matches are rejected for those key points for which the ratio of the nearest neighbour distance [3] to the second nearest neighbor distance is greater than 0.8. This discards many of the false matches arising from background clutter. The distance between two points x and y in Euclidean space generally means the shortest distance between the two points(x, y). If x and y are the histograms of two images, the Euclidean distance (ED) between two image is calculated by using the eqn (5)

$$ED = |x - y| = \sqrt{\sum_{i=1}^n |x_i - y_i|^2} \quad (5)$$

4. EXPERIMENTAL RESULTS ANALYSIS AND DISCUSSION

4.1. DATASET

There are various person image dataset publically available for person re-identification process. Every database contains different types of images with different configuration based on the camera settings and quality. Some database contains the image which has been taken from same location and some other contains the images taken from different locations. Based on the camera and location, the images are captured with the occurrence of various issues. For our experiment we have built the database by collecting the images from different benchmark dataset [1]. Our dataset contains two types of images which are indoor and outdoor. In total there are 240 images. All images are resized into 128*64 in pre-processing. The database details with different issues given in the Table I.

TABLE I. DATASET CONFIGURATION

Issues	Single person	Dataset Images (20 persons)
Occlusion	3	60
Illumination	3	60
Viewpoint	3	60
Pose Variation	3	60

The samples of outdoor images which are taken for the experiments from the database are figure out in the Fig 2. All the images are taken from the dataset which are from the public place surveillance camera.



Fig. 2 Sample Person Images of Indoor

4.2. EXPERIMENTAL RESULTS

Three experiments have conducted here for evaluate the performance of features for re-identification. The first experiment conducted is based on colour feature, second based on texture feature and the third based on shape feature.

For colour feature RGB has taken for texture SIFT and for shape canny edge features are taken. Finally all the experiments are analyzed using performance metrics.

The Fig.3 shows the sample results from the experiment. Input is the probe image and output is the similar image from gallery images .



Fig.3 Results on Various Issues

4.3. ANALYSIS AND DISCUSSION

4.3.1 PERFORMANCE METRICS

The performance of all the features evaluated based on different performance metrics. In existing research, rank based metrics only used for re-identification. In our experiment to analyze the performance in efficient way precision recall and accuracy of every feature for re-identification is evaluated.

Precision is defined as the average probability of relevant retrieval person images. The proportion of the true positives against all the positive results (both true positives and false_positives) gives the precision value. The expected identification accuracy is calculated using precision. It is calculated by using the eqn (6)

$$\text{Precision} = \frac{tp}{tp + fp} \quad (6)$$

Recall is the average probability of the complete retrieval of images. Recall in information retrieval is the fraction of the images that are relevant to the query that are successfully retrieved. It is computed by using the eqn (7)

$$\text{Recall} = \frac{tp}{tp + fn} \quad (7)$$

Accuracy is defines from the overall identification of the given input respect to the training images. This is the proportion of true results (both true positives and true negatives) among the total number of cases extracted. It is calculated by using the eqn (8)-

$$\text{Accuracy} = \frac{tp + tn}{tp + tn + fp + fn} \quad (8)$$

All the above performance metrics are calculated based on the tp (true positive), fp (false positive), tn (true negative) and fn (false negative). Tp is the correctly identified person images, tn is based correctly rejected persons, fp is mention as wrongly accepted person images and fn defines wrongly rejected persons.

4.3.2 PERFORMANCE ON RGB COLOUR FEATURE

To examine the performance of colour feature for person re-identification, the basic colour feature RGB is extracted from person's full body appearance. The colour feature mainly extracted from the clothing appearance of the person. The re-identification accuracy is evaluated using precision and recall from the experimental results. The re-identification accuracy of RGB is given in Table II.

TABLE II. Identification Accuracy Of RGB

Issues	Precision(%)	Recall (%)	Accuracy
Occlusion	66	56	62
Illumination	68	53	63
View point changes	72	62	71
Pose Variation	52	49	51

From the resulting Table II. We could observe that the RGB colour feature works poor for re-identification. It gives less identification accuracy and precision rate for all the issues. This feature works well only for the people who wear the same cloths without much variation in occlusion and illumination. Based the camera settings in different places also RGB colour feature gives various accuracy. So it is advised that the RGB is not suitable feature for identify the person.

4.3.3 PERFORMANCE OF SIFT TEXTURE FEATURE

SIFT is one of texture feature it covers all the pixels which has been occurred in all the objects. So it is considered as the basic level texture feature for person re-identification. It is extracted from the whole appearance of the person body. The key points of every part of input image are compared with the key points of the remaining images. The experiment has conducted on SIFT and the performance has evaluated and shown in Table III

Table III. Identification Accuracy of SIFT

Issues	Precision (%)	Recall (%)	Accuracy(%)
Occlusion	73	61	73
Illumination	50	61	63
View point changes	61	40	47
Pose Variation	51	61	60

From the resulting Table III. It is exemplified that the SIFT feature gives less accuracy for person re-identification for various issues. SIFT feature works better only for the persons who are appeared in the same place with same background without any changes in issues. So it works only for the objects without any movement. It does not suit well for person re-identification. Because the persons are captured from the movement of surveillance network. It is suggested that SIFT is considered to identify the people who are in same place with small changes.

4.3.4 PERFORMANCE OF CANNY EDGE BASED SHAPE FEATURE

Shape feature is extracted from different types of shapes of the person body. In this paper canny edge detection is used to detect the shape of the person based on the edge of the person body. It detects all the edges of the person body based on the given condition. Based on the edge detail the feature is extracted and the histogram of all the edges all the person is calculated . Finally all the histograms are stored in training dataset and the histogram of test image has compared with histograms of all the training images. This is also based on the full appearance of the person. The identification accuracy is shown in table 4.

Table IV. Identification Accuracy Of Canny Edge Based Shape Feature

Issues	Precision (%)	Recall (%)	Accuracy (%)
Occlusion	63	31	45
Illumination	30	63	65
View point changes	59	52	65
Pose Variation	32	62	60

The Table IV illustrate that the shape feature based on canny edge based shape context feature performs better for all the issues. This will detect the shape well even in variation of illumination pose variation. But it gives less accuracy for occlusion and view point changes. The reason is canny edge detect the edge of the occluded parts also. If the occlusion variation range is less the accuracy will not affected. Otherwise there will be more variation in accuracy.

From all the experiments we could analyze that every feature having its own properties it tries to overcome issues based on the properties. But it fails to overcome the other issues. All the features are not support to all issues.

5. CONCLUSION

In this paper the performance of three type of basic level feature for person re-identification is evaluated. All the experiments were conducted and analyzed effectively. The datasets of human are taken with different issues and analyzed. The Euclidean distance is used for matching the similarity between the images. From all the results and analysis it is inferred that the current features are not sufficient for the robust detection of person in surveillance. It gives very less accuracy for re-identification. By developing the new robust feature we can improve the re-identification accuracy well. The following extensions may be carried out for future research. New efficient feature can be proposed for re-identification. The new efficient feature extraction method can be developed for re-identification process.

REFERENCES

- [1] Vezzani, Roberto, and Rita Cucchiara. "Benchmarking for Person Re-identification." *Person Re-Identification*. Springer London, 2014. 333-349.
- [2] Bedagkar-Gala, Apurva, and Shishir K. Shah. "A survey of approaches and trends in person re-identification." *Image and Vision Computing* 32.4 (2014): 270-286
- [3] Zheng, Wei-Shi, Shaogang Gong, and Tao Xiang. "Reidentification by relative distance comparison." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35.3 (2013): 653-668
- [4] Zhao, Rui, Wanli Ouyang, and Xiaogang Wang. "Unsupervised salience learning for person Identification." *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013.
- [5] Ma, Bingpeng, Yu Su, and Frédéric Jurie. "Local descriptors encoded by fisher vectors for person re-identification." *Computer Vision–ECCV 2012. Workshops and Demonstrations*. Springer Berlin Heidelberg, 2012.
- [6] Martinel, Niki, and Christian Micheloni. "Re-identify people in wide area camera network." *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*. IEEE, 2012.
- [7] Liu, Chunxiao, et al. "Person re-identification: What features are important?." *Computer Vision–ECCV 2012. Workshops and Demonstrations*. Springer Berlin Heidelberg, 2012.
- [8] Bauml, Martin, and Rainer Stiefelhagen. "Evaluation of local features for person re-identification in image sequences." *Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on*. IEEE, 2011.
- [9] Cheng, Dong Seon, et al. "Custom Pictorial Structures for Re-identification." *BMVC*. Vol. 1. No. 2. 2011.
- [10] Doretto, Gianfranco, Thomas Sebastian, Peter Tu, and Jens Rittscher. "Appearance-based person reidentification in camera networks: problem overview and current approaches." *Journal of Ambient Intelligence and Humanized Computing* 2, no. 2 (2011): 127-151, Springer-Verlag 2011