International Journal of Advanced Information in Engineering Technology (IJAIET)
Vol.5, No.5, May 2018 ISSN: 2454-6933

A NOVEL APPROACH FOR PERSON RE-IDENTIFICATION USING BACKGROUND SUBSTRACTION AND SIFT FEATURES

SAINABA.K.SHAHUL,JOBI JOSE

Ilahia College of Engg, Ernakulam,Kerala
Email: sainabakshahul@gmail.com,jobijose03@gmail.com
Contact: +91-8606705447, +91-9496825891

Abstract: In video surveillance and realistic security, video based human re-identification plays a paramount aspect. Here, we proffer an eXtended Center-Symmetric version of local binary pattern (XCS-LBP) description for background modelling and subduction in videos and also uses scale invariant factor transform algorithm helps us to find out and extract shape difference regions and their location. Using those features we can re-identify person. Extensive researches are handle on two criterion public re-identification datasets including SDALF and SARC3D clearly shows that the intended schema for human re-identification using background subtraction and SIFT aspect fulfill the state of the art for video based re-identification significantly.

***Index Terms*:** Person Re-identification, Background Subtraction, SIFT

I. INTRODUCTION

In a computerized video surveillance, human re-identification plays a paramount task. It can be construe as comparing people over distinct camera visions. Due to changeover in visual appearances caused by differing conditions, re-identification can be contemplated as a challenging problem. In this endeavor, for human re-identification we propose the appearances and fluctuations from video framework of people. An utmost application in computer vision, the background subduction (BS) is the main steps being object trace and activity concession. Initialization and maintenance of Background model and Foreground disclosure are the main important tasks considered in background subduction mechanism. Different local texture descriptors have available for background modeling, mostly the local binary pattern (LBP) in behalf of its simplicity and quick computation.

Within the portrait pixel codes the LBP trait of a picture consists in constructing a histogram. For the sake of producing a long histogram, takes the first-order gradient information values of center pixel and neighboring pixel. In this papery, we proffer to prolong the variant by introducing a new neighboring pixel. The proffered XCS-LBP outturn its direct challenger for the background subtraction task. SIFT key points are first extricate from a set for reference picture and then saved in a database. Individually comparing each aspect from the recent image with the database, a key point of object is identified in an advanced image. Based on Euclidean distance for their aspect vectors candidate identical characters are identified.

II. SYSTEM ARCHITECTURE

In this papery, we proffer a novel accession for human re-identification using background subduction and SIFT factor thatt formulated the problem of person re-identification. Figure 1 layout the system architecture.

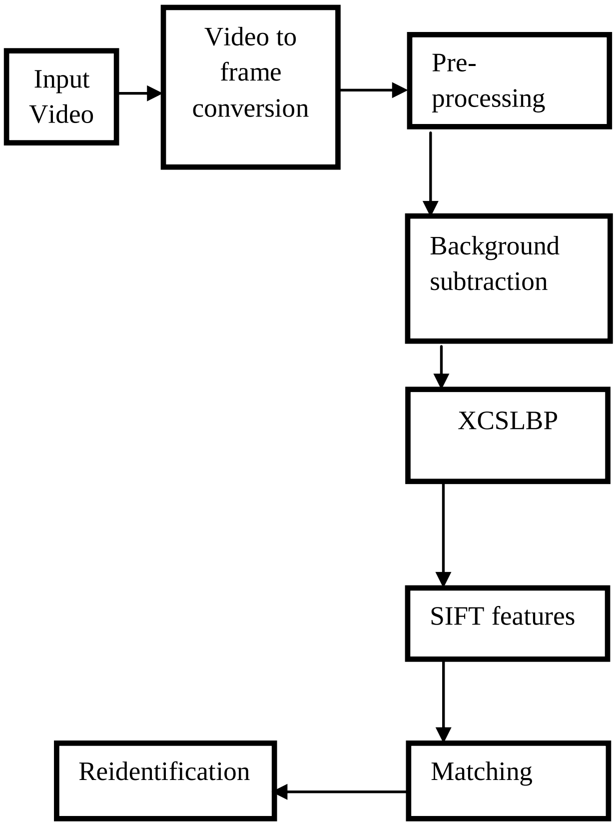

Figure 1: Block Diagram for Human Re-identification Using Background Subtraction and SIFT Features

For a given input video, we apply background modelling and subduction using XCS-LBP descriptor and also uses SIFT features learning person shape difference features. Ahead applies background subtraction and SIFT features, we lead to convert the video into different frames and perform preprocessing.

1. The XCS-LBP Descriptor

The original local binary pattern descriptor preface over a dominant local image descriptor. By thresholding the neighboring pixel value with the value of center pixel, it tags the image block pixels and binary numeral is the outcome. Prevailing portrait and the portrait representing the background model are concealing in the background subduction method. They become a texture-based version of the scene. Let a pixel at a certain location, contemplated as the center pixel c = (xc,yc) of a local neighbourhood composed of M equally spaced pixels on a circle of radius r . The LBP operator enforced to c can be disclose as:

$$LBP_{M,r}(c) = \sum_{i=0}^{M-1} s(g_i - g_c)\, 2^i \quad (1)$$

where gc is the gray value of the center pixel c and gi is the gray value of each neighboring pixel, and s is a thresholding function defined as:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

From (1), it is easy to show that the numeral of binary terms summed is $\sum_{i=0}^{M-1} 2^i = 2^M - 1$, the length of the result histogram is 2^M . Assuming an even number M of neighbouring pixels, the CS-LBP operator is delineate by:

$$CS\text{-}LBP_{M,r}(c) = \sum_{i=0}^{M/2-1} s(g_i - g_{i+(\frac{M}{2})})\, 2^i \quad (3)$$

where g_i and $g_{i+(\frac{M}{2})}$ are the gray values of centre symmetric cople of pixels, and s is the thresholding function defined as:

$$s(x) = \begin{cases} 1 & \text{if } x > T \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where T is a user-defined threshold. By construction, length of histogram resulting from the CS-LBP descriptor falls down to $1+\sum_{i=0}^{\frac{M}{2}-1} 2^i = 2^{M/2}$.

By compiling the gray values of a couple center symmetric pixels we proposed to prolong the CS-LBP operator. By considering the central pixel, the produced histogram is precise. The resulting descriptor lesser susceptible to buzz for the background subtraction application. The new LBP variant, called XCS-LBP (eXtended CS-LBP), expresses as:

$$XCS{-}LBP_{M,r}(c) = \sum_{i=0}^{M/2-1} s(g_1(i,c) - g_2(i,c))\, 2^i \quad (5)$$

where the threshold function s, which is used to determine the types of local pattern transition, is construed as a characteristic function:

$$s(x1 + x2) = \begin{cases} 1 & \text{if } (x1 + x2) \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

and where $g_1(i,c)$ and $g_2(i,c)$ are defined by:

$$\begin{cases} g_1(i,c) = (g_i - g_{i+(\frac{M}{2})}) + gc \\ g_2(i,c) = (g_i - g_c)(g_{i+(\frac{M}{2})} - g_c) \end{cases} \quad (7)$$

The intended XCS-LBP produces a shortened histogram than LBP, but it extracts further image minutiae than CS-LBP because it considers the gray value of the central pixel. The preferred descriptor appears to extra efficient for background modelling and subduction because it come out to be extra potent to buzz images than both LBP and CS-LBP.

2. SIFT Algorithm

a. Scale-Space Extrema Detection

We undertake by detecting key points in the SIFT algorithm. Key points are taken from maximal/minimal of the Difference of Gaussians (DoG) that materialize at conglomerate scales. The difference of succeeding Gaussian- dim images is taken after the image is convolved with Gaussian filters at various scales.DoG image $D(x,y,\sigma)$ is given by,

$$D(x,y,\sigma) = L(x,y,k_i\sigma) - L(x,y,k_j\sigma)$$

Where $L(x,y,k\sigma)$ is the convolution of the original image $I(x,y)$ with the Gaussian blur $G(x,y,k\sigma)$ at scale $k\sigma$,ie,

$$L(x,y,k\sigma) = G(x,y,k\sigma) * I(x,y)$$

Hence a DoG image between scales $k_i\sigma$ and $k_j\sigma$ is the difference of the Gaussian dim images at the same scales $k_i\sigma$ and $k_j\sigma$.

b. Key point Localization

Scale-space extrema disclosure outturn enormous numeral of key points, part of which are not stable. Interpolation of nearby data is assimilated to determine accurate position of applicant key point.Interpolation is executed by proving the quadratic Taylor expansion of the Difference-of-Gaussian scale-space function (x,y,σ) with origin as the applicant key point.

$$D(x)=D + (\partial D^T / \partial x)x) + (.5x^T ((\partial^2 D/\partial x^2)x)$$

where D and its derivatives are evaluated at the candidate key point and $x=(x,y,\sigma)^T$ is the offset from this point.To discard the low contrast key points,value of the second-order Taylor expansion D(x) is computed at offset $\hat{x}$. If the applicant key point is not powerful to miniature amounts of noise, the DoG function will have heavy responses along edges. We need to eradicate the key points that have poorly resolute locations but have high edge responses to increase the stability.

c. Orientation Assignment

This is the fundamental thread in achieving invariance to rotation as the key point descriptor. Key points can be illustrated relative to orientation and therefore achieve image rotation invariance. In order to perform all the computations at scale invariant manner first the Gaussian-smoothed image L(x,y,σ) at the keypoints scale is taken. For an image sample L(x,y) at scale σ,the gradient magnitude m(x,y) and orientation θ(x,y) are precomputed using pixel differences:

$$m(x,y)= \sqrt{(L(x+1,y)-L(x-1,y))^2 + (L(x,y+1)-L(x,y-1))^2}$$

$$\theta(x,y) = \text{atan2}\,(L(x,y+1) - L(x,y-1)\,,\, L(x+1,y) - L(x-1,y))$$

d. Key point Descriptor

Locate key points at peculiar scales and orientation assignments are performed at the previous steps. This causes invariance to image location, scale and rotation.

EXPERIMENTAL RESULTS

In this paper for surveillance and forensic in 3D environments we prefer a new dataset, namely **3DPeS (3D People Surveillance Dataset)**.It specifically designed for re-identification tasks, and relevant to countless tasks such as human disclosure, tracking, action analysis and trajectory analysis.

1. Performance Evaluation

To appraise the enforcement of whole step including segmentation, tracking, shadow analysis, trajectory and activity concession the 3DPeS dataset can be used. In this case, the decent enforcement assessment metric **cumulative matching characteristic (**CMC**)** curve is, exhibit how performances advances as the numeral of requested images increases. The CMC curve represents the expectation of finding the correct match in the top n matches and is analogous to the ROC curve for detection problems. If the CMC curve for the matching function is given, the probability that any of the M best matches is correct generates the synthetic recognition rate (SRR) as follows:

$$SRR(M) = CMC(N/M), \qquad (8)$$

where CMC(k) is the rank k recognition rate and M is the numeral of pedestrians in the scene taken from a large testing dataset of size N. Figure 2 shows the moderated CMC curves of both methods.

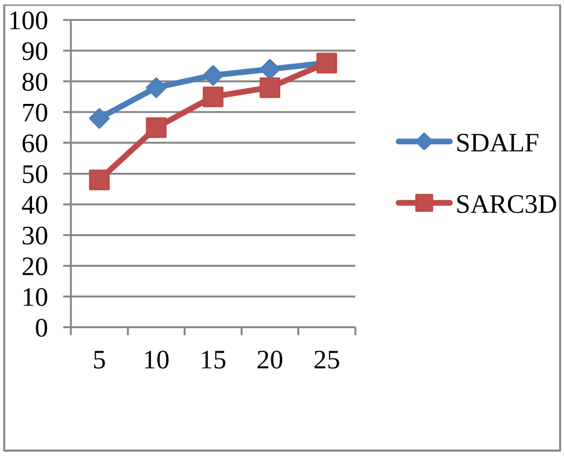

Figure 2: Results of Our Preliminary Tests Using SDALF and the SARC3D Method

In this papery, we use two public benchmarks SDALF, SARC3D to evaluate our proposed background subduction and SIFT model.

- SDALF (Symmetry Driven Accumulation of Local Features) proposed in 2010 by Farenzzeena et al [4].
- SARC3D a multi-view method established on a 3D body model we proposed in 2011 [5].

The first is a purely two dimensional method. It subsists in the eradication of features. It exemplars three correspondent aspects of the human appearance including the overall chromatic content, the dimensional arrangement of colours into reliable zone, and the existence of recurrent regional concept with high entropy. We compared the above method with 3D-based method SARC3D. Humans are disclosure and traced in each graded camera with their appearances, location and orientation extricate and used to place, scale and orientate a simplified 3D body model. Fig. 3 shows some frames of the SARC3D method in action.

The AD HOC system was used to track peoples and extract their features in both cases. A few selected appearance images were extracted from each video progression and used for the re-identification by using the same algorithm. We selected randomly an individual portrait from one by one video progression in the case of SDALF. But in the case of SARC3D model 3 and 5 portraits for each progression were randomly selected. For each method we performed 10 test runs using sequences of 100 people randomly selected from the dataset. Figure 4 shows some selected test queries. As the graph shows, by using multiple views and 3D data the re-identification performances significantly increase, especially if the number of returned ranked matches is limited. For instance at rank-1 SDALF returns 26.78% of correct matches, while SARC3D 37.51%. At rank-5 46.42% for SDALF against 67.8% for SARC3D

Figure 3: Frames of SARC3D Method

(a)

(b)

Figure 4: Example queries to our re-identification database. (a) Probe image (for SARC3D this is just one of the images used for the model creation). (b) Top 10 results (sorted left to right). First row SDALF results, second row SARC3D results. The correct match is highlighted in green.

CONCLUSION

In this endeavor, we proffer a novel accession for human re-identification employing background subtraction and SIFT features method addressing video human re-identification problem. We propose an eXtended centre symmetric local binary pattern (XCS-LBP) description for background modelling and subduction in video and SIFT features learning persons shape difference features. XCS-LBP combines the fortitude of the authentic local binary pattern (LBP) and the centre symmetric (CS) LBPs. Thus the new variant XCS-LBP (eXtended CS-LBP) outturn a curtailed histogram than LBP, by its CS-construction. The proposed XCS-LBP descriptor qualitatively and quantitatively outperforms the other descriptors, making it a genuine applicant for the background subduction undertaking in computerized applications. The scale-invariant feature transform (SIFT) is an algorithm in computer vision to detect and describe local traits in images. SIFT features helps us to find out and extract shape difference regions and their location. Using those features we can re-identify person. We organized extensive research on two civics available video based person re-identification datasets to substantiate our process. Experimental results demonstrated that our model outperforms other state of the art methods in most cases and verified that our background subtraction and SIFT feature exemplary for human re-identification is beneficial for the understanding certainity in person matching.

REFERENCES

[1] Ojjala, T., Pieetikainen, M., and Maennpaa T. (2002).Multiresolution gray-scale and rotation invariant texture classifications of local binary pattern. IEEE Trans. on Pattern and Machine Analysis, pages 971–987.

[2] Xue, H, Song, L., Suun, J., and Wue, M. (2011). Hybrid centre-symmetric local pattern for dynamic background subtraction. In IEEE Int. Confe. on Multimedian and Expo, pages 1–6.

[3] Haoe Liue, Jiaashi Femg, Meiein Qi, Jianguoe Jiamg and Shuichaeng Yan, Fellow,End to end comparative attention network for person reidentification, IEEE TRANSACTION ON IMAGE PROCESSING, VOL. 14, NO. 8, JUNE 2016.

[4]M. Farenzzena, L. Buzzani, A. Peerina, V. g, and M. Cristayi. Person re-identification by symetry-driven accumulations on local feature. In Proce. of CVPR, pages 2360–2367. IEEE, June 2010.

[5]D. Baltiery, R. Vezzani, and R. Cucchiara. Sarc3d: a new 3d body model for people tracking and re-identication. In Proce. of ICIAP, Ravennea, Italy, Sept. 2011.