Vol 8. No. 2 Issue 2 – August, 2015

African Journal of Computing & ICT

www.ajocict.net

Human-Objects Re-identification by Low-Level Features and Mid-level Attributes in A Multi-Camera Surveillance Environment

M. E. Irhebhude[1], F. Ajakaiye[2], A.E. Evwiekpaefe[3] & H.I.K. Etsu-Ndagi[4]
Faculty of Science,
Department of Computer Science
Nigerian Defence Academy
Kaduna, Nigeria.
Emails: [1]hyelmart4lyf@gmail.com, [2]a.fiyinfoluwa@gmail.com, [3]contact_abraham@yahoo.com, halisku25@yahoo.com.

ABSTRACT

A novel technique for people re-identification is proposed in this research article based on using low-level colour features and mid-level attributes. The low-level colour histogram bin values were normalised to 0 and 1. A publicly available dataset (VIPeR) and a self constructed dataset have been used in the experiments conducted with 7 clothing attributes and low-level colour histogram features. These 7 attributes were detected using features extracted from 5 different regions of a detected human object using an SVM classifier. The low-level colour features were extracted from the regions of a detected human object. These 5 regions are obtained by human object segmentation and subsequent body part sub-division. People are re-identified by computing the Euclidean distance between a probe and the gallery image sets. We have shown that the proposed 7 mid-level attributes and the low-level features results in improved performance accuracy for people re-identification.

Keywords: People re-identification, People tracking, Surveillance, Occlusion, Recognition

African Journal of Computing & ICT Reference Format:
M. E. Irhebhude, F. Ajakaiye & A.E. Evwiekpaefe (2015): Human-Objects Re-identification by Low-Level Features and Mid-level Attributes in A Multi-Camera Surveillance Environment. Afr J. of Comp & ICTs. Vol 8, No. 2, Issue 2. Pp 141-152.

1. INTRODUCTION

A fundamental task for a distributed multi-camera surveil- lance system is to recognise individuals in diverse scenes obtained using two or more cameras at different times and locations. Person re-identification is a long term people surveillance and monitoring task, where individuals or a group of people are differentiated from several possible targets in diverse scenes, obtained from different cameras distributed over a network of locations of substantial dis- tances, in the presence of occlusions, difference in view angles, lighting conditions and time. In a surveillance scenario, an individual disappearing from a particular camera view needs be matched with similar human objects present in one or more other views obtained at different physical locations, over a period of time, and be differentiated from numerous other human objects in the same views.

In a typical surveillance / video monitoring task, it can help to find out if a particular individual who enters and exits a building is the same person identified within another different building; within a public space, work environment, university campus, school, train station, airports etc. The views of surveillance footage may be taken from different, angles and distances, backgrounds, lighting conditions and various degrees of occlusions. Although in general, significant feature variations could be present in a significant variety of clothes worn by people, vast majority of public may choose to wear ordinary clothes with similar appearance in daily living. Such characteristics which bear a mid-level semantic meaning can be exploited for a person re-identification task. In this research article, we will consider mid-level semantic attributes as valued vari- ables for the person re-identification problem. For example, we will consider the trouser to either be coloured or bright.

We propose a selective parts-based approach for low- level feature representation of a pedestrian and for mid-level feature attribute detection for human description. This approach helps to reduce misalignment, avoidance of the background and helps in clothing attributes detection, which help improve re-identification accuracy. A specifically captured dataset alongside existing publicly available dataset; Viewpoint Invariant Pedestrian Recogni- tion (VIPeR) were used in the experiments conducted. For clarity of presentation this article is divided into a number of sections as follows: immediately following this section is section 2, which examines the state-of-art works in people re-identification.

Section 3 explains the proposed method for person re-identification. Section 4 describes how the parts of a holistic human figure were detected to enable detailed clothing attribute detection. Section 5 shows us the list of clothing attributes used for the proposed person re-identification task. Section 6 gives us the results of the various experiments performed. Section 7 presents exper- imental analysis with their respective performance results, while section 8 and 9 gives the conclusion and future work.

Vol 8. No. 2 Issue 2 – August, 2015
African Journal of Computing & ICT

www.ajocict.net

2. LITERATURE SURVEY

The goal of object re-identification is to correctly identify all instances of the same visual object at any time or location [1]; meaning, choosing the most probable object among sets of possible matches of consecutive observations of the same target at different camera views [2]. In [3] three features were accumulated; entire colour content, colour regions, texture characteristics of recurrent region to form Symmetry Driven Accumulation of Local Features (SDALF) and used on three datasets to give a novel state-of-art performance in object recognition and re-identification In [4], authors combined [3]'s SDALF technique with mid-level semantic feature attribute to identify candidate objects. Further the importance of attributes and how relevant attribute features can be selected for object re-identification task was also demonstrated. Random forest technique was used by [5] to determine the importance of individual feature attributes under different circumstances of various roles for object classification.

A framework, Multiple Component Matching (MCM) was proposed in [6] for object re-identification. MCM was explained as an ordered set of sequences con- taining several components with simulated parts generated to cater for illumination variation. Authors however established that simulated components increased the computation com- plexity. To correct the computation complexity issue authors vectorised and clustered the MCM to form a prototype. The matching were done in the dissimilarity space with text information used as a query for image retrieval. Mean Rie- mannian Covariance Grid (MRCG) in [7], modeled clothing information to describe the human object for recognition. Covariance matrix was used to describe images of fixed sizes with equal grid structures and averaged to get the Mean Riemannian Covariance (MRC) that describes the object for re-identification.

In [8] HOG features were trained to detect body parts; top, torso, leg, left arm, right arm. Covariance's of colour gradient and orientation was computed on each region including the full body to get discriminative signature used for people re-identification. In [9] the standard LBP was modified by setting dimensionality at 16 to form the Simplified LBP (SLBP) to detect people's head and face. In order to re-identify people; authors used [7]'s MRCG technique to model the detected head and face so as to capture a discriminative signature. An optimised Speed Up Robust Feature (SURF) named Camellia Key Point was used in [10] to describe grayscale (to eliminate variation in colours) candidate objects and used for re-identification on CAVIAR datasets with the threshold set at 15. In [1] colour samples were modelled using fuzzy K-Nearest Neighbour (KNN) algorithm to segment candidate objects into eleven culture colours. Probability Colour Histogram (PCH) plot were used to identify an object at a set threshold after comparing two targets in intra and inter camera scenar- ios. People in a crowded environment can be identified by integrating appearance features: selective upper body patch and candidate position and direction of travel using a landmark-based model [11]. Analysis showed that the proposed technique performed better than the full body based integration. In [12], SURF features was proposed for interest point extraction using Sum of Quadratic Difference (SQD) as a point correlation tool for object identification in a distributed camera network.

In a similar scenario, [13] proposed an unsupervised iterative brightness transfer func- tion (BTF), a technique to handle the variability that occurs in illumination conditions. BTP helps to map brightness values between intra camera views while cumulative BTF helps to adapt colours in inter camera views for people re- identification. In a low quality camera network; [2] used a Colour Structure Descriptor (CSD) by extracting dominant colours from regions of interest (shirt and pant); derived CSD by evaluating the differences of dominant colours between the two targets and proposed a so-called Target Colour Structure (TCS) for people re-identification. A two feature approach was proposed in [14] for object recognition, i.e., Haar and Dominant Colour Descriptor (DCD) features. Haar features of the foreground mask recognised an object in the first technique while DCD works by partitioning the detected foreground object into two, then using the dominant colours of both regions as descriptors for object recognition.

In [15] two techniques were proposed; Red Green Blue (RGB) colours were used alongside the height feature his- togram and transformed normalised RGB colours plus the height feature histogram techniques to identify objects using histogram matching. Instead of recognising objects using a distance measure, [16] proposed Ensemble RankSVM for ranking image sets with the correct match having the highest ranking score. A comparison between rank and distance measure techniques for object re-identification was conducted. Ensemble RankSVM was however recommended because of the scalability of the technique. In [17], ULBP and Hue Saturation Value (HSV) histogram were used as features extracted from body segmented into 3 parts of a detected target to capture local texture and colour features.

These features alongside direction of view captured different identifiers for 3 views; front, back and side that helped in person re-identification. In [18], a model which is a function of pose was developed to capture human appearance. With the rectified pose prior image specific person's feature of colour and textures were extracted; re-identification and identification of targets became more robust to viewpoint on the trained dataset. In [19], persons were re-identified by accumulating local weight map histogram features from 3 areas of a segmented human body. The local weighted histograms were trained for optimal weight map. These local weight map histograms were integrated to form a feature vector used for identification.

Vol 8. No. 2 Issue 2 – August, 2015
African Journal of Computing & ICT

www.ajocict.net

In [20], the use of middle-level clothing attribute information was described to assist in person re-identification. Re-identification perfor- mance was improved by treating clothing attributes as real value variables. In their pre-processing steps, a body part-based representation approach was proposed by extracting HSV colour histogram and HOG as features. A further contribution was the generation of a large-scale dataset that contains more samples and camera views than the currently available public datasets. In [21], people were re-identified by a combination of features; hue, saturation histogram and Saliency Maps from selected body parts. In [22], a technique that identifies human action and appearance based on colour and optical flow models was proposed. The mean features from two regions of a detected candidate identified a person's action and appearance. The colour features were extracted from 8 colour spaces; R, G, B, H, S, Y, Cb, Cr channels respectively.

In [23], instead of solving people identification problem using ranking and distance measures, Takač et. al. used an appearance based learning algorithm such as SVM and the Naive Bayes classifiers to identify people. Finally, [24] proposed a mid-level identification approach called the Optimised Attribute Re-identification. 21 attributes were proposed and detected using SVM. As reported by [24], concentrating errors, biasness, match- ing errors and human surveillance costs, has given rise to the need for the automation of re-identification tasks. Despite the past and present efforts to solve the automation of the re-identification problem using various techniques [25], it still remains a research area, where much research effort are needed, due to the fact that conventional biometrics such as face recognition has failed as a result of insufficient region of interest (ROI) detail for extracting robust features. Further, in exploiting other visual features such as ap- pearance of a person, most features used in literature have not been sufficiently discriminative enough for low quality inter-camera differentiation, due to changes in a person's appearance, differences in view angles, changes in lighting conditions, presence of background clutter and occlusion etc [25].

3. PROPOSED RE-IDENTIFICATION FRAMEWORK

This section presents the operational details of the pro- posed human object re-identification system. The process of re-identifying a person in a video surveillance system generally includes three broad steps: human object detection; feature capture and representation and object classification (see figure 1).

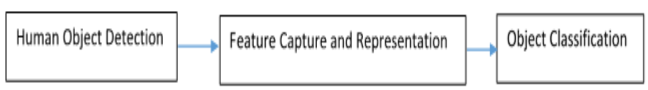


Figure 1: Human re-identification process

Figure 2 illustrates the detailed block diagram of the proposed person re-identification system. Sections 4 6 presents the underlying algorithmic details of each of the functional blocks of the figure 2 below. Fundamentally, in the proposed system, the re- identification of a person is carried out by jointly making use of so-called low-level features of a person's appearance (i.e. a detailed colour histogram of central body part regions, see section 4-A) and so-called mid-level features captured from a person's head, torso and leg regions (e.g. dark head, coloured shirt, dark trouser etc). More specifically the low-level feature representation of person's appearance is defined by detailed colour histograms which are normalised and obtained in regions of an initial body part segmentation and a subsequent sub-division (see section 4-A); while the mid-level feature representation of a person's appearance is defined by a higher-level description of the same regions that determines for example whether the shirt/trouser is dark/coloured or not and head is dark or not etc. The details of these functional blocks can be described in the following section.

4. HUMAN BODY PART-BASED FEATUREREPRESENTATION

Prior to the detection and analysis of a human body parts or segments for subsequent feature extraction, the full human body needs to be detected in a scene. For this purpose we utilised the object detection technique of [26] which uses HOG features for human localisation. Once the full body is identified as defined within a single rectangle, a body part segmentation and a subsequent sub-division is carried out. Finally the a detailed feature analysis is carried out (see section 5 and 6) within the above regions that is finally used for person re-identification.

A. *Body region segmentation and sub-division*

Assuming a standing and upright human, body region segmentation and sub-division helps subsequent capture of specific features of a segmented human object. This seg- mentation is performed by splitting the rectangular region containing the complete human figure into three parts, namely; head, torso, and leg (see figure 2). Further sub division of these three regions into smaller regions of interest (ROIs) is done by further splitting the; head region into three horizontally separated, equally sized sub-regions, the torso and leg regions are divided into equally sized, 3 × 3

Vol 8. No. 2 Issue 2 – August, 2015
African Journal of Computing & ICT

www.ajocict.net

rectangular sub-regions, as depicted by figure 2. In order to minimise the effects of consideration of the background regions in further analysis, only the middle rectangular patch is selected from the head region and the four middle patches, placed vertically, are selected from the torso and the leg regions, for subsequent capture of low-level colour histogram features and further attribute selection.

B. Low-level feature extraction and representation

The next step after body regions segmentation and sub- division is the colour histogram based feature detection and representation of the five centrally spaced regions. For each of the five said regions a so-called RGB 3D-8 bin colour histogram is extracted by (see figure 2) dividing each colour channel (i.e. R,G and B) into 8 bins and concatenating into a single feature vector of length $8 \times 8 \times 8 = 512$. Consequently, the appearance of a person is described by a feature vector, obtained by concatenating features of the five centrally located patches; giving a total feature length of 2560 (see figure 3).

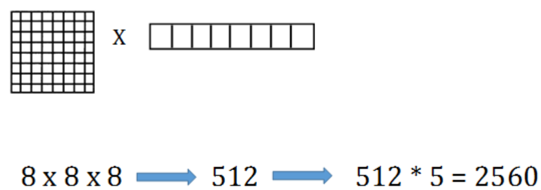


Figure 3: Low-level feature concatenation

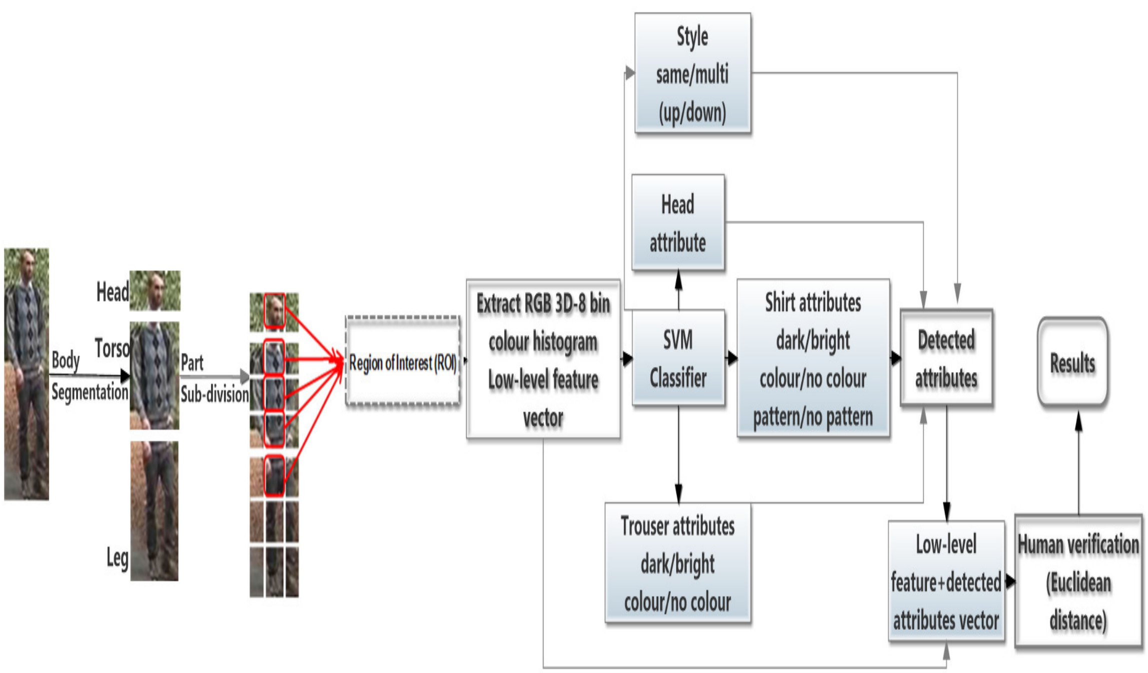


Figure 2: Proposed system for person re-identification

5. CLOTHING ATTRIBUTE REPRESENTATION

Aimed at creating a more detailed representation of a human figure by adding further higher level features to the low-level feature descriptor obtained above (see section 4-B) the said five regions are further analysed to determine seven attributes that determines a higher-level appearance of the human body. Figure 4 illustrates the seven attributes defined. One attribute is defined from the head-region, namely the 'head- colour'. Three attributes are defined from the shirt re-gion, namely the 'shirt-colour', 'shirt-brightness' and 'shirt- pattern', Two attributes are defined from the trouser region, namely, 'trouser-colour' and 'trouser-brightness'. Finally, one attribute is defined for describing the overall appearance, namely, 'clothing-style'. Each of the above attributes can take two possible values as tabulated in table I. Hence the value of each of the attributes can be represented by a binary number 1, or 0, for e.g. dark-shirt with 1 and non-dark shirt with 0.

Vol 8. No. 2 Issue 2 – August, 2015
African Journal of Computing & ICT

www.ajocict.net

A. Clothing attribute value determination

The medium-level attribute values of test human objects were determined by using a Support Vector Machine (SVM) classifier to train on hand annotated attributes with known values from known sample regions of a training image dataset (see section 7-A). As a result of the above each detected human figure's medium-level features will be represented by a seven ele- ment vector with each element being either a zero or a one.

B. The combined feature vector

Figure 5 illustrates the combined feature vector that com- prises of the low-level 3D-8 bin colour histogram features and medium level features that are represented by the above mentioned attributes. This combined feature vector defines the detected human and will subsequently be used in human re-identification.

RGB 3D 8 bin Colour histogram features	Described attributes

Figure 5: Total feature length

6. EXPERIMENTS

Two datasets were used for experiments, a self-captured set of new content and the most popular database used by other researchers, i.e., VIPeR.

A. ***Self-captured dataset***

The captured database has 118 frames which comprises of footage relevant to 6 different people taken from two different cameras. All images are scaled to a size of 128×48 pixels. In our experiments the cameras are named as A and B and the set of images captured by Cam B are used as the gallery images and the set of images captured by the Cam A are used as the probe image set. The performance of the proposed algorithm for person re-identification is evaluated by matching each test image in Cam A against the images in Cam B, the gallery image set. Figure 6 shows some examples of the detected persons in the self-constructed dataset. This dataset contains predominantly indoor images with challenges in illumination changes due to changes in artificial lighting within the building.

Figure 6: Samples from the self-captured data set

B. ***The VIPeR dataset***

The VIPeR dataset contains 632 pedestrian image pairs captured by two cameras having different viewpoint, pose and lighting. Images are scaled to size 128×48 pixels. In our experiments we name the two camera as Cam A and Cam B. In the experiments conducted the set of images captured by the Cam B are considered the gallery set and those captured by the Cam A are considered as the probe image set. The algorithmic performance is evaluated by matching each test image in Cam A against the Cam B gallery. Some selected example images from the VIPeR dataset are illustrated in figure 7

Vol 8. No. 2 Issue 2 – August, 2015
African Journal of Computing & ICT

www.ajocict.net

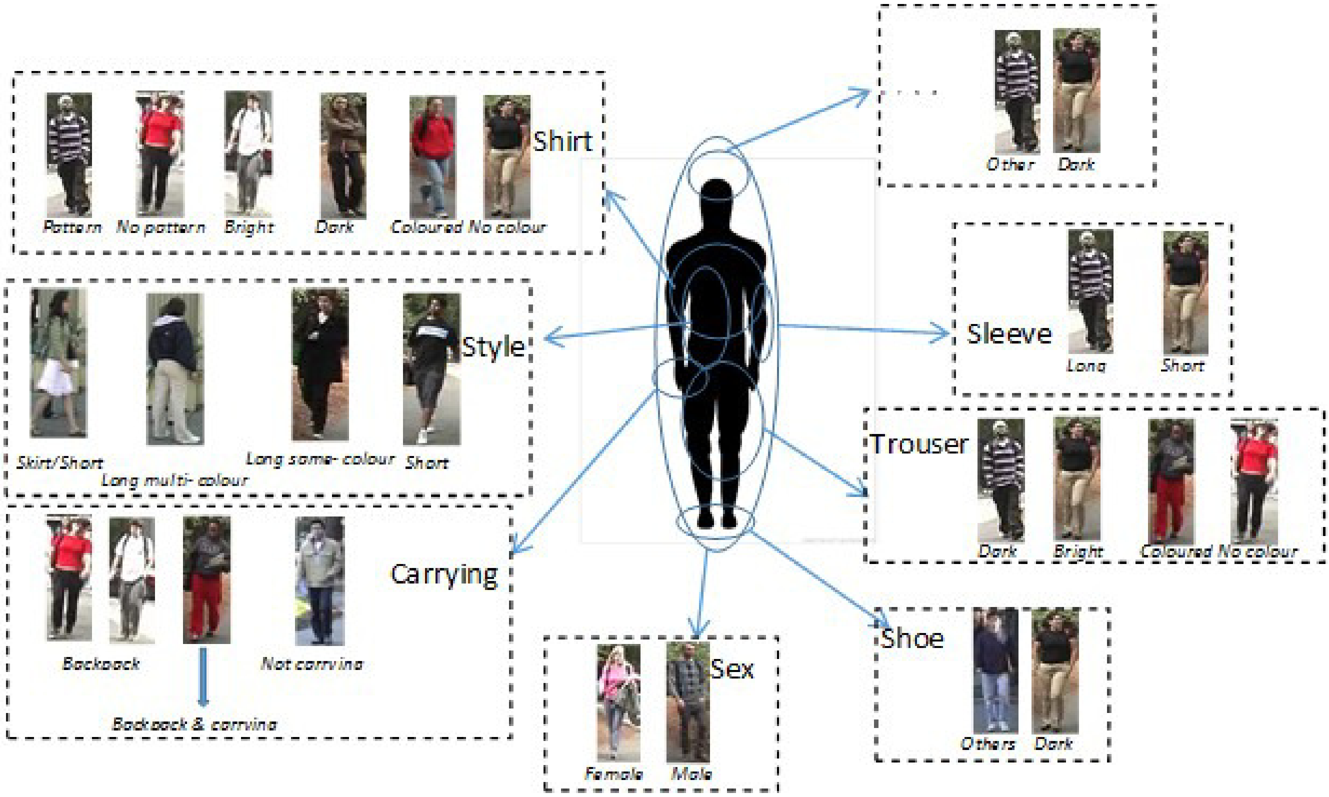


Figure 4: Definition of medium-level attributes

Table I: Attributes description and values

Number	Attributes	Value1	Value2
1	Shirt-Colour	Coloured	No Colour
2	Shirt-Brightness	Bright	Dark
3	Shirt-Pattern	Patterned	No Pattern
4	Clothing-Style	Single colour up/down	Multi-colour up/down
5	Head-Colour	Dark	Other
6	Trouser-Brightness	Dark	Bright
7	Trouser-Colour	Coloured	No No colour

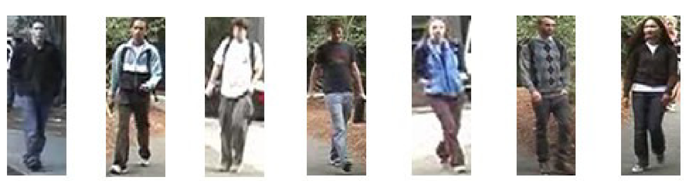


Figure 7: VIPeR data samples

C. Evaluation and metrics used

The database used for evaluation be it the VIPeR dataset or the self-captured dataset is first divided into two sets, i.e., the training image set and test image set. Approximately half of the images are used for training and the remaining half is used for testing. We train an SVM classifier on both the training and validation portions, while re-identification performance is reported on the held out test portion. A person from the query image set is re-identified using a distance metric between itself and each of the candidate images in the gallery image set. The low-level, distance measure, dL, between a query image, Iq and a candidate image from the gallery image set Ig is defined as follows:

Vol 8. No. 2 Issue 2 – August, 2015
African Journal of Computing & ICT

www.ajocict.net

$$d^L(I_q,I_g) = \sum_l d_l^L(L_l(I_q),L_l(I_g)) \quad (1)$$

where Ll (Iq) and Ll (Ig), refers to the extracted type l low- level features from the query and gallery images i.e Iq and Ig respectively and dL is the corresponding distance measure for the feature type l.
For the clothing attributes, the distance measure is defined as follows:

$$d^A(I_q,I_g) = \sum_a d_a^A(A_a(I_q),A_a(I_g)) \quad \textbf{(2)}$$

where Aa(Iq) and Aa(Ig) are the attribute encoding ’a’ of the query image Iq and the candidate gallery image Ig
Given the above definitions, the Euclidean distance metric between a query image and a gallery image based on the low-level features is defined as follows:

$$d^L = \sqrt{\sum_i (q(l_i|x_{q,i}) - g(l_i|x_{g,i}))^2} \quad (3)$$

where li|xq,i refers to the it h low-level feature of the query image given all other features of the query image and li|xg,i refers to the it h low-level feature of the gallery image given all other features of the gallery image. Similarly, the Euclidean distance metric between the query image and a gallery image based on the attribute-space is defined as follows:

$$d^A = \sqrt{\sum_i (q(a_i|x_{q,i}) - g(a_i|x_{g,i}))^2} \quad (4)$$

where all terms can be defined in a manner similar to that defined in equation 3.

In literature, the standard performance evaluation metrics used in person re-identification are matching performance at rank n, cumulative matching characteristic (CMC) curves, and normalised Area Under the CMC Curve (nAUC) [24]. The matching performance at rank n reports the probability that the correct match occurs within the first n ranked results from the gallery image set. This is obtained by calculating the Euclidean distances between a query image and all images in the gallery image set and ordering the matches in ascending order of matching error. The match with the smallest error is considered the rank-1 image and so on. The CMC curve plots the recognition for all rank values, n, and the nAUC summarises the area under the CMC curve (Note: the ideal nAUC is 1.0 and nAUC of 0.5 defines match obtains simply by ’chance’). However, the measures used for the performance eval- uation of the proposed person re-identification algorithm are limited to the rank score illustrated by the associated cumulative matching characteristic (CMC) curves.

7. EXPERIMENTAL RESULTS AND ANALYSIS

This section presents the experimental results and a de- tailed analysis. The performance of the proposed approach was considered using three different matching metric mea- sures namely, a) matching based on low-level features only
b) matching based on medium-level attribute signatures only and c) matching based on both low level features and attributes, combined.

A. Attributes detection

After the extraction of low-level colour features they can be used in the colour based recognition of values of the seven attributes of a human figure defined in Table I. The VIPeR database was used for the attribute training and testing. From the images captured for Camera A, each attribute value was manually annotated. The manually annotated information from Camera A, for a given attribute (say for e.g. shirt- colour) was used in training an SVM. The testing was done on images captured by Camera B. Each attributes value was determined using the relevant trained SVM. This training and testing processes were carried out for each attribute, separately, using a different SVM. Table II records the detection accuracies obtained for each of the attributes. The highest accuracy has been obtained for ’Style’ and the lowest accuracy has been recorded for the Head region in deterring whether it is dark or not. The latter is due to the high possibility of presence of individuals with darker skin tone and these individuals getting mixed up with people who are turning the back of their head to the camera.
The average accuracy for the detected attributes is 77.9%.

B. Matching performance analysis

Figure 8 illustrates the CMC curves when low-level features and attributes are used for the representation of detected people, both as individual metrics and together, i.e. as a combined metric. When the combined feature set is used the figure 9 illustrates the same graphs plotted within the narrow range of Rank-1 to Rank-20. The results indicate that up to Rank-5 the combined feature set performs better than the individual feature sets.

Vol 8. No. 2 Issue 2 – August, 2015
African Journal of Computing & ICT

www.ajocict.net

However above Rank-5 a better accuracy of recognition is demonstrated when using the Attributes only. This indicates that the detailed low level colour histogram features add details to the person's Attributes making the matching more accurate at up to Rank-5. However the use of low-level colour features only is not recommended due to relatively poor performance. A detailed study revealed that the low- level colour features although providing details for higher ranked matches, when used independently varies signifi- cantly between images of even the same person. Having the Attributes considered in addition allows the combined features to more accurately define an object.

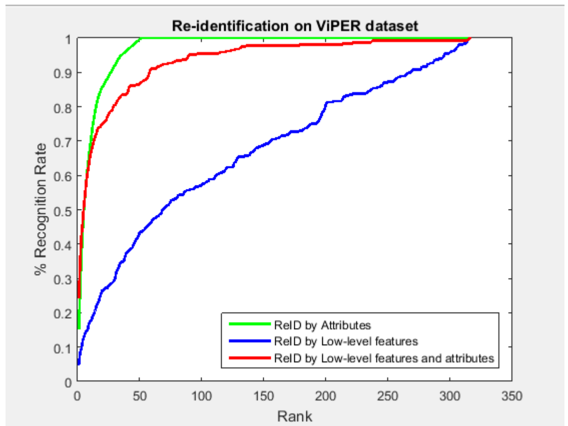


Figure 8: Cumulative matching characteristic curves of pro- posed technique

Table II: Attributes classification accuracies based on VIPeR dataset

Number	**Attributes**	**Value2**	**Detection accu-racy**
1	Shirt-Colour	No Colour	79.4%
2	Shirt-Brightness	Dark	73.4%
3	Shirt-Pattern	No Pattern	87.8%
4	Clothing-Style	colour Multi-colour up/down	90.7%
5	Head-Colour	Other	66.5%
6	Trouser-Brightness	Bright	70.9%
7	Trouser-Colour	No No colour	76.4%
	Mean		**77.9%**

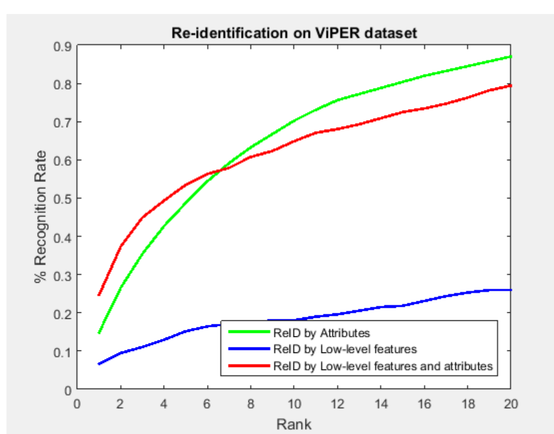


Figure 9: Cumulative matching characteristic curves of pro- posed technique plotted within the narrow range of Rank-1 to Rank-20.

Vol 8. No. 2 Issue 2 – August, 2015
African Journal of Computing & ICT

www.ajocict.net

The average accuracy obtained by averaging over all Rank's was 62%, 97.6% and 92.1% respectively for low- level features, attributes and their combination. Table III compared the performance of the proposed ap- proach to that of the method proposed in [24] that proposed a low-level feature based approach dependent on colour and textures for initial attribute detection and an subsequent attribute only based approach for person re-identification. The results have been tabulated for the same set of training and test images obtained from the VIPeR image database.

The results tabulated in Table III show that the at Rank-5 and above the proposed approach when only the Attributes are used and the combined set of Attributes and low-level features are used performed significantly better than the method proposed in [24] a method popularly used as a benchmarking algorithm in literature. However at Rank-1 the proposed method when only the Attributes are used performs less accurately as compared to the benchmark algorithms. It is noted that the benchmark algorithm of [24] is based on a larger (hence more detailed set of medium-level features) set of attributes (21 attributes) as compared to the number of attributes used by the proposed technique (7 attributes). This is the likely reason for it to perform better than the proposed algorithm at Rank-1 when only the Attributes are use.

However when the combined low-level colour features and medium-level Attributes are used the proposed algorithm works better. This is due to the additional detail of the objects definition included by the low-level colour attributes that are used in the proposed approach.

The proposed low-level feature set only includes colour features from the RGB representation of the image. However the low-level feature set that the algorithm in [24] uses for attribute detection uses both colour features and texture features. The colour features, show less in number is spread across three different representations of object colour (RGB, HS and YCbCr). Our detailed investigation revealed that when colour features of the same object when represented in different colour features are used, a significant amount of redundant information is used in the training process. This affects the accuracy. Further global texture features are very much subjected to changes due to background clutter, over/under exposed images etc, that could also affect in a negative manner if texture features are also used alongside colour features. Figure 10 illustrates bar graphs comparing the perfor- mance at different Rank scores. Results in Table III also tabulates the performance of the proposed approach when combined features are used and the self-captured dataset with more challenging images are used for experimentation

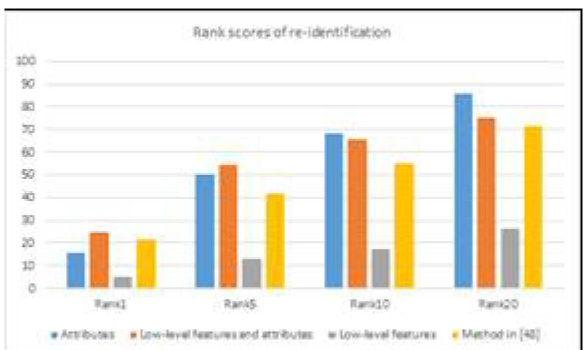


Figure 10: Rank scores re-identification performance

Figure 11 illustrates an example of matching gallery images for a probe image from the query image dataset when using the VIPeR dataset (top row) and the self-captured (bottom row) dataset. It is seen that the query image matches with a number of candidates from the gallery image database where the person has turned with respect to the camera angle of view.

Table III: Person re-identification accuracy

VIPeR	Rank1	Rank5	Rank10	Rank20
Attributes	15.5	50	68.4	85.8
Low-level features and attributes	24.7	54.4	65.5	75
Low-level features	5.1	13	17.4	26.6
Method in [24]	21.4	41.5	55.2	71.5
Self-constructed				
Proposed technique	5	35.6	56	74.6

Vol 8. No. 2 Issue 2 – August, 2015

African Journal of Computing & ICT

www.ajocict.net

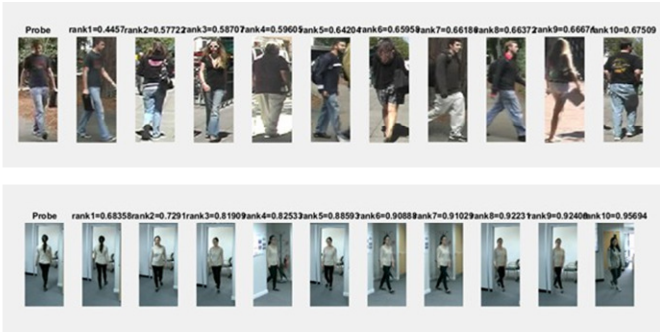


Figure 11: Human re-identification on both datasets

The above results indicate the superior performance obtainable from the proposed approach.

8. CONCLUSION

Finally, we have shown that detailed colour featured captured in known localities of a human figure in the form of a 3D colour histogram with a finite number of bins can be used to accurately determine attributes of a human body that can then be used together with the low-level colour features for person re-identification. Accuracy figures of approximately 75% and 85% have been obtained when using combined Attribute and low-level features and Attributes only, respectively at a rank of Rank-20.

9. FUTURE WORKS

In literature the performance of people re-identification systems have always been demonstrated and evaluated on still images. The possibility of implementing the proposed technique within a real-time video analytic scenario so as to demonstrate the applicability of this system in a real world system, is proposed as future work. It was also revealed that the mid-level attributes detection performance could benefit from some performance improvements. Investigation of the use of additional features, the use of more effective feature reduction techniques and feature combinations are recommended. Further investigating the use of effective feature weighting, based on training data in obtaining the combined feature vector for representing an human object is also recommended

ACKNOWLEDGEMENTS
This work was completed with the support of Nigerian Defence Academy, Kaduna, through the Tertiary Education Trust Fund (TETFUND) intervention, Nigeria.

Vol 8. No. 2 Issue 2 – August, 2015
African Journal of Computing & ICT

www.ajocict.net

REFERENCES

[1] Angela D'Angelo and Jean-Luc Dugelay. People re-identification in camera networks based on probabilistic color histograms. In IS&T/SPIE Electronic Imaging, pages 78820K–78820K. International Society for Op- tics and Photonics, 2011.

[2] Federica Battisti, Marco Carli, Giovanna Farinella, and Alessandro Neri. Target re-identification in low-quality camera networks. In IS&T/SPIE Electronic Imaging, pages 865502–865502. International Society for Optics and Photonics, 2013.

[3] Michela Farenzena, Loris Bazzani, Alessandro Perina, Vittorio Murino, and Marco Cristani. Person re- identification by symmetry-driven accumulation of lo- cal features. In Computer Vision and Pattern Recogni- tion (CVPR), 2010 IEEE Conference on, pages 2360– 2367. IEEE, 2010.

[4] Shaogang Gong Ryan Layne, Timothy Hospedales.Person re-identification by attributes, 2012.

[5] Chunxiao Liu, Shaogang Gong, ChenChange Loy, and Xinggang Lin. Person Re-identification: What Features Are Important?, volume 7583, book section 39, pages 391–401. Springer Berlin Heidelberg, 2012. http://dx.doi.org/10.1007/978-3-642-33863-2\ 39.

[6] Riccardo Satta, Giorgio Fumera, Fabio Roli, Marco Cristani, and Vittorio Murino. A multiple component matching framework for person re-identification, 2011.

[7] Slawomir Bak, Etienne Corvee, Francois Bremond, and Monique Thonnat. Multiple-shot human re-identification by mean riemannian covariance grid, 2011.

[8] Slawomir Bak, Etienne Corvee, Francois Bremond, and Monique Thonnat. Person re-identification using spatial covariance regions of human body parts, 2010.

[9] Francois Bremond Etienne Corvee, Slawomir Bak. People detection and re-identification for multi surveil- lance cameras, 2012.

[10] Omar Hamdoun, Fabien Moutarde, Bogdan Stanci- ulescu, and Bruno Steux. Person re-identification in multi-camera system by signature based on interest point descriptors collected on short video sequences, 2008.

[11] Andrea Cavallaro Riccardo Mazzon, Syed Fahad Tahir. Person re-identification in crowd. Pattern Recognition Letter, 2012.

[12] Icaro Oliveira de Oliveira and Jose Luiz de Souza Pio. People reidentification in a distributed camera network, 2009.

[13] Clemens Siebler, Keni Bernardin, and Rainer Stiefel- hagen. Adaptive color transformation for person re- identification in camera networks, 2010.

[14] Francois Bremond Monique Thonnat Slawomir Bak, Etienne Corvee. Person re-identification using haar-based and dcd-based signature, 2010.

[15] Henri Bouma, Sander Borsboom, Richard J. M. den Hollander, Sander H. Landsmeer, and Marcel Worring. Re-identification of persons in multi-camera surveil- lance under varying viewpoints and illumination. pages 83590Q–83590Q, 2012. 10.1117/12.918576.

[16] Shaogang Gong Tao Xiang Bryan Prosser, Wei-Shi Zheng. Person re-identification by support vector ranking, 2010.

[17] Kevin Krucki, Vijayan Asari, Christoph Borel-Donohue, and David Bunker. Human re-identification in multi-camera systems. In Applied Imagery Pattern Recognition Workshop (AIPR), 2014 IEEE, pages 1–7. IEEE, 2014.

[18] Ziyan Wu, Yang Li, and Richard J Radke. View-point invariant human re-identification in camera net- works using pose priors and subject-discriminative features. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 37(5):1095–1108, 2015.

[19] Tetsu Matsukawa, Toshiya Okabe, and Yuuki Sato. Per- son re-identification via discriminative accumulation of local features. In Pattern Recognition (ICPR), 2014 22nd International Conference on, pages 3975–3980. IEEE, 2014.

[20] Aoxue Li, Luoqi Liu, Kangping Wang, Siyuan Liu, and Shuo Yan. Clothing attributes assisted person re- identification. 2014.

[21] Sebastian Hommel, Darius Malysiak, and Uwe Hand- mann. Efficient people re-identification based on mod- els of human clothes. In Computational Intelligence and Informatics (CINTI), 2014 IEEE 15th International Symposium on, pages 137–142. IEEE, 2014.

[22] David Gero´nimo and Hedvig Kjellstrom. Unsupervised surveillance video retrieval based on human action and appearance. In Pattern Recognition (ICPR), 2014 22nd International Conference on, pages 4630–4635. IEEE, 2014.

[23] Boris Takac, Andreu Catala, Matthias Rauterberg, and Wei Chen. People identification for domestic non-overlapping rgb-d camera networks. In Multi-Conference on Systems, Signals & Devices (SSD), 2014 11th International, pages 1–6. IEEE, 2014.

[24] Ryan Layne, Timothy M Hospedales, and Shaogang Gong. Attributes-based re-identification. In Person Re- Identification, pages 93–117. Springer, 2014.

[25] Shaogang Gong, Marco Cristani, Chen Change Loy, and Timothy M Hospedales. The re-identification challenge. In Person Re-Identification, pages 1–20. Springer, 2014.

[26] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1, pages 886–893. IEEE, 2005.

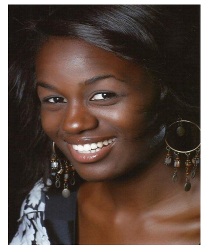
Vol 8. No. 2 Issue 2 – August, 2015
African Journal of Computing & ICT

www.ajocict.net

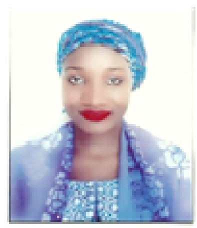
Authors' Short Biographies

Martins E. Irhebhude obtained his tertiary and master degree education in Edo State, Nigeria in 2003 and 2008 respectively. He is concluded his PhD research degree in 2015 with the Computer Science Department in Loughborough Uni- versity, UK under the supervision of Eran A. Edirisinghe PhD, a Professor of Digital Image Processing. Martins is a staff of the Nigerian Defence Academy, Kaduna State, Nigeria since 2004 and currently engages in teaching and research activities in and around the Defence Academy. Research interests includes: object detection, people tracking, people re- identification, object recognition and vision related researches.

Abraham E. Evwiekpaefe Abraham is a lecturer in the Department of Computer Science at the Nigerian Defence Academy, Nigeria. He holds a B.Sc and MSc degrees in Computer Science. He is currently a PhD research student. Interest areas include: Software Engineering, Mobile Comput- ing and Object detection.



Fiyinfoluwa Ajakaiye is a Lecturer in the De- partment of Computer Science at the Nigerian De- fence Academy, Nigeria. Current research inter- ests include: Human Computer Interface, Object tracking and Surveillance. Fiyinfoluwa obtained a BSc in Software Engineering and MSc in Com- puter Animation. She is currently in pursuit of a PhD.



Halima Ibrahim Kure Etsu-Ndagi is an Assistant Lecturer in the department. Of Computer Science, Nigerian Defence Academy, Kaduna. She obtained MSc in Information Security and Computer Forensics - University of East London, UK 2012, BSc Software Engineering University of East London, She can be reached at halisku25@yahoo.com or by phine on +2348038904883