**Paper Review Report:**

## People silhouette extraction from people detection bounding boxes in images, Christophe Coniglio, Cyril Meurie, Olivier Lézoray, Marion Berbineau, [Pattern Recognition Letters](https://www.sciencedirect.com/science/journal/01678655)

[Volume 93](https://www.sciencedirect.com/science/journal/01678655/93/supp/C), 1 July 2017, Pages 182-191, Citation: <https://doi.org/10.1016/j.patrec.2016.12.014>

**1. Introduction**

This paper introduces the concept of people detection (PD), noting that it is a key element in many human action recognition applications. It presents the advances in PD, beginning with handcrafted techniques utilising support vector machines (SVMs) and then onto recent methods employing convolutional neural networks (CNNs). The problems of PD methods applied to action recognition are explored. The most important of which is the imprecise information provided by bounding boxes used in PD methods for people segmentation (PS) to locate objects. Furthermore, it notes that little research has been done for PD methods used for PS in images. Therefore, this paper proposes a new strategy for precise PS using bounding boxes as still images obtained from PD methods.

**2. Proposed Method**

The proposed method has three stages. The first is a pre-treatment stage for the images of the bounding boxes. The second stage is the probability map estimation stage which estimates the conditional probability of every pixel in a bounding box image as either a foreground or background pixel. The last stage is the segmentation refinement which uses the probabilities estimated in the previous stage to classify the pixels as either foreground or background pixels to obtain the people silhouette. Various numbers of PS techniques are considered for each of the various stages. Thus, in order to determine the best configuration of the techniques for each stage as well as the parameters for those techniques, a genetic algorithm is employed.

**2.1. Pre-treatment**

There are three types of pre-treatment techniques. They include a colour space change, filtering and a colour invariant transformation. The colour space change has the advantage of allowing better differentiation between colours. Filtering helps to reduce noise and the colour invariant transformation reduces unwanted light effects like high brightness.

**2.2. Probability map estimation**

There are three types of probability map estimation techniques. They include shape template, colour histograms and saliency map priors. There are three different shape template priors and two different colour histogram priors. The three shape template priors include: the mean shape template prior, the multiple shape template prior and histograms of oriented features combined with SVMs. The two types of colour histogram priors are: colour histograms and colour histogram strips. Colour histograms are computed over the whole bounding box, while colour histogram strips concatenate histograms computed on strips of the bounding box. The advantage these techniques over other techniques like gaussian mixture models is that they are much faster to compute. Saliency map priors are also considered, as the person in the bounding box is a salient object.

**2.3. Segmentation refinement**

There are two types of segmentation refinement techniques considered in this paper. The first is graph cut segmentation which creates people silhouettes by optimizing a discrete energy function via computing the minimum cut of the graph associated with the image. The second is super-pixel graphs and in this paper the two super pixel techniques considered are those in [1] and [2].

**2.4. The genetic algorithm**

A genetic algorithm is employed to select the techniques to use in each stage and to tune their parameters. The algorithm is applied on a population of 120 different chromosomes (where chromosomes represent proposed methods) made of genes. A gene can either encode the use of a particular technique in a stage or it can set the parameters of the technique. The genetic algorithm comprises four steps. The first is an initialization step which constructs a list of candidate methods by randomly initializing the genes of the chromosomes, causing techniques and parameters to be randomly selected in each stage. The second is a crossover step which randomly selects two parent chromosomes, and generates a child chromosome by randomly selecting a gene from each pair of matching genes from those two parent chromosomes. The third is the mutation step where a gene or a set of genes from the child chromosome is altered according to a mutation rate. The last is the selection step, where each chromosome is ranked according to a fitness score and then the bottom 50% are discarded. This algorithm is iterated until the best chromosome has not changed for 10 generations.

**2.5. Novelty of the proposed method**

The novelty of the proposed method is seen through the use of a genetic algorithm which obtains the best set of methods that can be used for PS in still images. This is verified by the little work conducted on PS in still images. Only a handful like the method proposed by Jojic et al [3] have been conducted, but they do not use a genetic algorithm.

**3. Results from experimentation**

Once the genetic algorithm was conducted, eight different methods/solutions were proposed. Each have varying numbers of techniques and parameter settings for the 3 stages. The datasets used to test the proposed methods were the VIPeR [4], PRID 2011 [5], INRIA Person [6] and BOSS European [7] datasets.The proposed methods were compared against two state-of-the-art methods using the same datasets. The state-of-the-art methods were the methods from Migniot et al [8] and Yang et al [9]. When the proposed methods were compared against the state-of-the-art methods, the method that came out on top was the graph-cut segmentation scheme. This scheme outperformed all the other proposed methods as well as the state-of-the-art methods with an average score of 0.884. Furthermore, six out of the eight proposed methods outperformed both state-of-the-art methods and all eight proposed methods outperformed the method from Yang et al. However, one disadvantage these proposed methods had was that they were slower than the best state-of-the-art method, namely the one from Migniot et al.

**4. Conclusion / Future works**

In conclusion, it seems that the proposed methods which were derived from the genetic algorithm were indeed superior to state-of-the-art methods. Based on this it seems safe to suggest that future work on this method would be beneficial in obtaining better techniques for still image PS. When analysing the results, this paper suggested that future work to tackle the problem of people re-identification from their extracted silhouettes ought to be conducted. Furthermore, it seems that further work conducted on improving the processing times of these methods should be explored.

**References:**

[1]  R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, S. Süsstrunk, SLIC superpixels compared to state-of-the-art superpixel methods, IEEE Trans. Pattern Anal. 34 (11) (2012) 2274–2282.

[2]  P. Buyssens, I. Gardin, S. Ruan, A. Elmoataz, Eikonal-based region growing for efficient clustering, Image Vision Comput. 32 (12) (2014) 1045–1054.

[3] N. Jojic, A. Perina, M. Cristani, V. Murino, B. Frey, Stel component analy- sis: modeling spatial correlations in image class structure, in: CVPR, 2009, pp. 2044–2051.

[4]  S. Gong, M. Cristanio, S. Yan, C. Loy, Person Re-Identification, Springer, 2014.

[5]  M. Hirzer, C. Beleznai, P.M. Roth, H. Bischof, Person re-identification by descriptive and discriminative classification, in: Image Analysis, Springer, 2011, pp. 91–102.

[6]  N. Dalal, B. Triggs, INRIA Person Dataset, Technical Report, INRIA, 2005.

[7]  Boss, Boss European project (on board wireless secured video surveillance), 2009. URL: http://www.multitel.be/image/research- development/ research- projects/boss.php.

[8] C. Migniot, P. Bertolino, J-M.Chassery, Automatic people segmentation with a template-driven graph cut, in: ICIP, 2011, pp. 3149–3152.

[9] C. Yang, L. Zhang, H. Lu, X. Ruan, M.-H. Yang, Saliency detection via graph-based manifold ranking, in: CVPR, 2013, pp. 3166–3173.