RESEARCH ARTICLE OPEN ACCESS

Effect of Noise on Likelihood Neighbor Extraction of Foreground Images

Sabna A B[1], Majo John[2], Diana Davis[3]

Asst Professors, Jyothi engg college

Trissur, Kerala

India.

ABSTRACT

Extraction of portion of image has so many applications in the field of image processing and its computer application. The hard application needs high resolution cut whereas soft applications can have any from many. Extraction of image may not be the as such application users may want, but it is the prior step in many computer related image processing application. The automatic and probability based method give awesome result in normalized graphical cut. The result improves when add similarity of a set associative pixels along with probability while making cut. The result varies with number of accounting neighboring pixels

Keywords- Pixel Similarity, Segmentation, Foreground extraction Normalized Graph cut approach; Automatic segmentation.

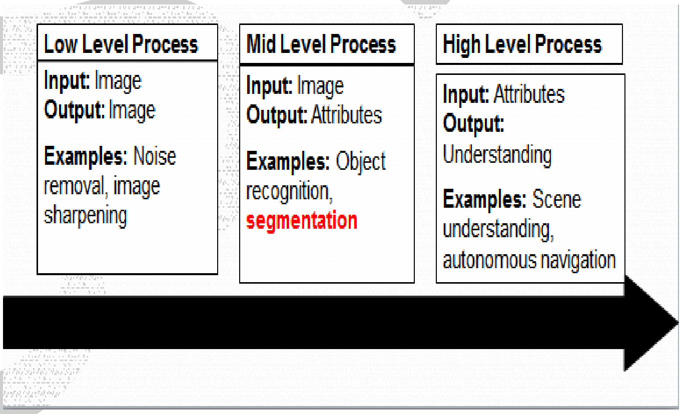
I. INTRODUCTION

Segmentation is one of the most important problems in image processing and analysis. It is the process of extraction of a particular portion (foreground) of the rest of the image (background). The segmentation is used as an initial step in many other image processing operations. Some of the applications depend heavily on the initial models obtained as the result of the segmentation. The accuracy and the original result of that application depends on the initial silhoughty obtained. Making the initial models perfect is a problem and challenge in all aspect.

Multi-view segmentation is nothing but segmentation applied on different views of the same object obtained from different angles. It has application in animation, 3D modeling, video creation etc. There are a lot verity methods exist for segmentation; like region based, quad tree based probability based, similarity based etc and all these are comes under two main category of segmentation, pixel based and volume strip based segmentation. All methods can be used to do normal segmentation and also the multiview segmentation.

A. Operational level of segmentation.

Image processing can be divided into three operational levels. Low level processes like noise removal, image sharpening etc. Midlevel processes include object recognition, segmentation etc whereas the third level include autonomous navigation and scene understanding etc. The sequence of for-levels forms prerequisite for the later one.



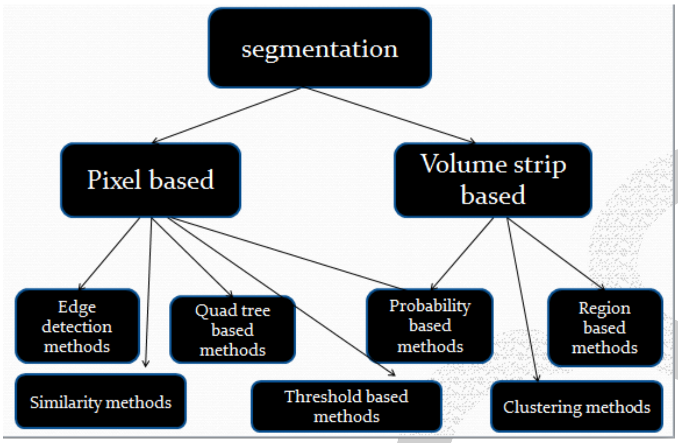
B. Classical defenition and goals.

Classically, image extraction-segmentation is defined as the partitioning of an image into non overlapping, constituent regions which are homogeneous with respect to some characteristic such as intensity, color or texture.

The ideal goal of segmentation is to identify the semantically meaningful components of an image and thus to simplify or change the representation of an image into more meaningful and easier to analyses. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions results after segmentation are significantly different with respect to the same characteristic.

II. TYPES OF SEGMENTATION

There are various types of image segmentations [1], [2]. Figure 1.1 names some important ones. Edge based segmentation, probability based image segmentation, similarity based image segmentation, graph cut based image segmentation, quad tree method, segmentation based on function variation, region growing method, region shrinking method are some examples



Edge detection methods: Edge based segmentation is also known as contour based image segmentation. Edge detection significantly reduces the amount of data and filters out useless information [7], while preserving the important structural properties in an image. In this type the objects are selected by means of edges.

Edge detection refers to the process of identifying and locating sharp discontinuities in an image. The discontinuities are abrupt changes in pixel intensity which characterize boundaries of objects in a scene. Edges are detected using a combination of operators for intensity gradient, texture discontinuities and color variation etc. Supplementary processing step must follow to combine edges in to edge chains. Edge chains are used in the segmentation process. Canny's edge detector is a good example.

Clustering methods: Clustering methods [8] are usually iterative methods to partition an image into a number of clusters or groups. The initial clusters have to be re order to make it include only similar characteristics at the end of the operations. K-means clustering is an example clustering method. The quality of the solution depends on the initial set of clusters and the value of K.

Thresholding methods: Thresholding is the simplest way of doing image segmentation. It is a technique which bands some range of values and permits other range. The key parameter in the thresholding process is the selection of threshold value [9]. If a single threshold value is used, it is called a global thresholding. Adaptive thresholding uses different threshold values for different parts of the image. A task well suited for local adaptive thresholding is segmenting text from the image. During the thresholding process, individual pixels in an image are marked as "object" pixels if its value is greater than the threshold and as "background" pixels otherwise.

Region based methods: Region based segmentation is used for image characteristics to map individual pixels in an image to sets of pixels called regions that might correspond to an object or a meaningful part of one. The various techniques comes under region based methods are local techniques, global techniques, splitting techniques and merging techniques. The effectiveness of region growing algorithms depends on the application area and the input image. If the image is sufficiently simple, local techniques can be effective. However, on difficult scenes, even the most sophisticated techniques may not produce a satisfactory segmentation. Hybrid techniques using a mix of the methods above are also popular.

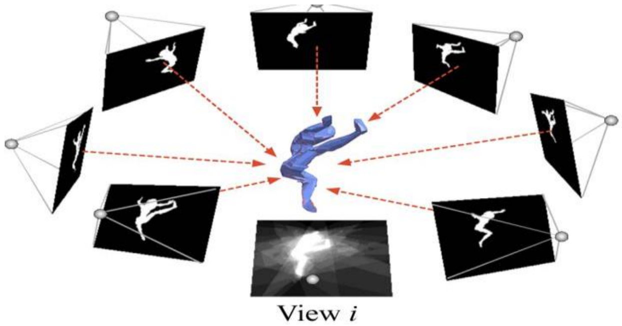
Quad tree methods: This method is also known as region splitting and merging method of segmentation. Quad tree method [10] considers image to be segmented as a tree. The root of the tree represents the whole image. If the image at root is found non uniform (not homogeneous), then it is split into four son-squares (the splitting process). This process continues recursively until no further splits or merges are possible or all sub trees are homogeneous. Conversely, if any child squares are homogeneous, they can be merged (the merging process). The QT structure allows dividing an image within a complete tree representation, including neighboring information. This spatial information can be further used by a merging strategy which joins the QT leaves using color and edge information.

Bayesian methods: Bayesian methods are also known as probabilistic model of image segmentation. This method calculates the maximum a posteriori probability [11] of each pixel. This involves the use of inferred probabilities rather than observed probabilities. Hence it involves the use of conditional probability as in any other case of Bayesian model segmentation. Section II will give a detailed review of Bayesian method.

A. What is multiview segmentation?

Multiview segmentation is a new area of segmentation and is found to be very useful in computer vision applications. In its simple form, multiview is nothing but multiple views of a single object from different angles. Multiview segmentation extracts object from each view. Combining different angle pictures of an object from different angles can easily generate a

3 dimensional voxel of the object. Even though Multiview segmentation is mainly used to generate the 3D form of an object from its 2D image, it has many other computer vision applications like object tracking, motion censoring, images and video synthesis, animation etc.



The concept of multiview is clearly depicted in the above picture. One object is viewed from different angles and thus we get different images. One obvious thing that can understand from the figure is that there will be something in common between these images like the background color, foreground color, some pixel values etc.

Multiview can be segmented out with any type of above discussed segmentation method. But Bayesian method seems to be more appropriate for multiview segmentation.

III. BAYESIAN METHOD-AN OVERVIEW

Bayesian method of segmentation classifies pixels into groups based on the posterior probability of them. Posterior probability explains the inferred probability rather than observed probability. In the context of images, the observed probability or simply the probability of a pixel implies that the probability of a pixel to be assigned to a specific intensity or color, whereas the inferred probability implies the probability of a particular pixel to have a specific intensity or color given that its neighboring pixel is having a certain observed value. From the observed value of pixels Bayesian method infer the value of the other pixels. This will yield more accurate values.

From Bayesian perspective, there are known values and unknown values [12]. The known quantity is the data, denoted as D. The unknown quantities are the parameters (eg. mean, variance, or missing data), denoted byθ. Then to make inferences about the unknown quantities, Bayesian stipulate a joint probability function that describes the unknown quantities behave in conjunction with known data, p (θ, D) [13].

Bayesian Theorem for a given parameter θ can be represented as

$$p(\theta \mid D) = p(D \mid \theta) \quad p(\theta) / p(D)$$

or

Posterior $\propto$ likelihood x prior

Where,

The prior is the probability of the unknown parameter and represents what was thought before seeing the data.

The likelihood is the probability of the data given prior parameter and represents the data now available.

The posterior represents what is thought given both prior information and the data just seen.

The likelihood and the posterior value is not variable ones and they are the observed quantities. The only quantity through which we can improve posterior probability is by using the prior knowledge. As much knowledge we have prior to operation that much efficiency will be there in the posterior probability.

A. *Why Bayesian Model?*

All the other segmentation methods treats multiview segmentation as multiple number of individual segmentation. Bayesian aspect of multiview segmentation treats the different views as a correlated one and utilizes previous segmentation details for further segmentations. This is accomplished by using conditional probability. The optimization is done by means of graph cut mechanism. Normal graph cut has a problem that there is a chance that the isolated edges get cut by minimum cut maximum flow theorem. This can be avoided if implemented with normalized graph cut.

Bayesian method involves conditional probabilistic calculation of events. If any other method is used to do silloughty estimation of segmentation, it does not make any change from multiple numbers of 'single' view segmentations. But Bayesian method efficiently uses the previous segmentation details (observed quantities) to be in cooperated with the current segmentation (inferred quantity). This will make current segmentation more efficient, because the views are of the same object from different angles.

Bayesian aspect of Inferred probabilities (from both prior and observed values) and the dependencies among them are clear from the given dependency graph

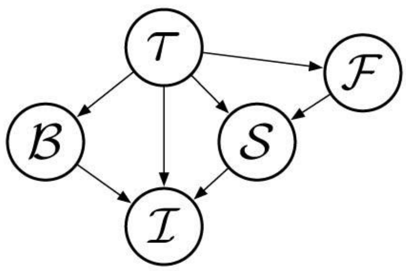


Fig 2. 1 Dependency graph of an image *I*

Here,
τ represents the prior knowledge;

β is the background color model;

S is the binary silhouette map and

F is the foreground spatial model.

As shown in figure 2.1, the dependency graph, an image observation *I* is influenced by the background color at the corresponding pixel location of background β and by whether the background is occluded or not at that location of Silloughty *S*, which is itself governed by the projection of the foreground region *F*. It is assumed that *F* and *β* to be independent, which can be argued because of shadows cast by the foreground, can change the background appearance. It is assumed that shadows have a negligible impact on the background colors.

Bayesian model utilizes this dependency to efficiently estimate the silloughty of each image in the multiview .Thus it provide better convergence than other methods of segmentation.

From Dependency graph, we can form the joint probability equation as

$$Pr(S, F, \beta, \tau) = Pr(\tau)\ Pr(\beta/\tau)\ Pr(F/\tau)\ Pr(S/ F, \tau)\ Pr(I /\beta, S, \tau) \quad \ldots\ldots\ldots(1)$$

where,
Pr (τ), Pr ((β/τ), Pr (F/τ) are the prior probabilities of the scene, the foreground, and the background, respectively.

Pr(S / F, τ) is the silhouette likelihood that determines how likely is a silhouette given the foreground shape. Since F is unknown, this term can be approximated by a spatial consistency term that determines how likely a silhouette S_i is given all silhouettes $S^{i \neq j}$ is and it measures the probability that a pixel belongs to the foreground region from the other views.

Pr(I /β, S, τ) is the image likelihood term that models the relationship between the image observations, colors and the background information which measures the similarity.

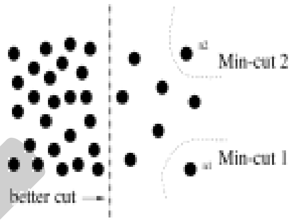
IV. MIXING SIMILARITY WITH PROBABILITY USING GRAPH CUT

Graph cut mechanism sees the image as a graph with set of edges and vertices to implement pixels and its relationship or similarity with other pixels [19]. Graph cut works by maximum flow minimum cut criteria.

Graph has to be partitioned into foreground region and background region based on the nature of the graph vertices. The partition can be done by simply removing edges connecting the two parts. The degree of dissimilarity between these two pieces can be computed as total weight of the edges that have been removed. In graph theoretic language, it is called the cut. The mathematical representation is given as

$$cut(A;B) = \sum w(u; v),\ u \in A,\ v \in B \ldots\ldots\ldots\ldots(2)$$

The minimum cut criteria occasionally favours cutting isolated nodes in the graph. An example case is shown here



Let the edge weights are inversely proportional to the distance between the two nodes, then the nodes which lies apart have a very small weight. Any cut that partitions out individual nodes on the right half will have smaller cut value than the cut that partitions the nodes into the left and right halves. So the minimum cut happens at less weighted edges and of course, that happens to separate an isolated node. To avoid this unnatural bias of partitioning out small sets of points, normalized graph cut can be used
.

IV. 8WAY NEIGHBORHOOD NORMALIZED GRAPH CUT

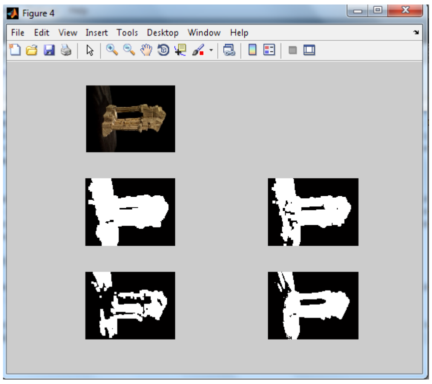
Associative neighborhood cut results better because it takes into account both similarity within groups and dissimilarity across groups along with conditional probability. 8 way Normalized cut computes the cut cost as a fraction of the total edge connections to all the nodes in the graph. The concept of Normalized cut [20] can be extended as

$$Ncut(A;B) = \sum(w(u; v)) / \sum(w(u; V)) \ldots\ldots\ldots(3)$$

Where u∈A, v∈B and V is the vertex set.

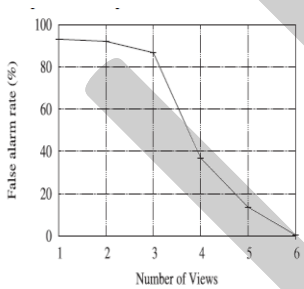
The normalized cut produces better results because it takes into account both similarity within groups as well as dissimilarity across groups along with conditional probability

The example output of temple data set [17] segmentation is given

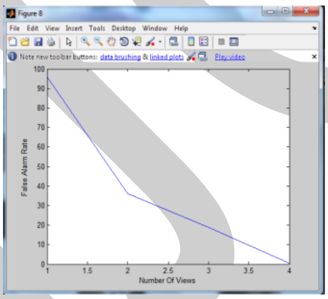


The normalization is done by checking a pixels 8 neighbors. A pixel has 8 neighbors, 4 diagonal, 2 raw and 2 column.. If a pixel is found to change its background status, it checks the status of its neighboring points. In silhouette, the pixel having 0 represent background and the pixel having a value 1 is called foreground. If 4 or more neighboring pixel has a foreground status, then only the current pixel is expected to change its value from 0 to 1. That is nearly 40 percentage or more neighboring pixel with foreground status allows to set a pixel's foreground status.

Method is evaluated with Dino data set [14], temple data [17] set and Kungfu girl data set [21]. The evaluation graph which measures the false alarm rate (the rate at which the pixels are falsely identified) is shown for both normal graph cut and normalized cut. Normalized graph cut makes convergence even faster than normal procedure.



Graph cut (a)



(b) Normalized cut

V. THE EFFECT ON NOISE OF N-WAY ASSOCIATIVE CUT

In the set associative neighborhood cut, the noise effect will be varying according to the number of associative pixel. If we choose the upper bound of associative pixel, ie 8 pixel associative, then a pixel is selected as foreground pixel only if all its vertical, horizontal, and diagonal pixels are similar. So the chance of occurring noise is less .On the other hand of lower bound , 1 pixel associative, then the pixel is selected as foreground pixel if any one of its its vertical horizontal, or diagonal pixels is similar. So the chance of occurring noise is high. From this it's clear that the chance of noise pixel inversely proportional to the value of n in n way associative neighborhood cut. Noise pixel decreases with increase I the walue of n.

$$\text{NOISE PIXEL} \;\infty\; 1/n,$$

Where n is the number of associate neighborhood pixel

A. *Related work.*

Kalin kolev et. al, in [3] implements a probabilistic formulation of multiview segmentation, given a series of images from calibrated cameras. This is a silhouette based approach. Instead of segmenting each image separately in order to build a 3D surface consistent with these segmentations, they had computed the most probable surface that gives rise to the images. They used two separate Bayesian frame work which are independent of each other. One for background voxel and one for foreground voxel. Voxel is nothing but small volume areas. This framework converts the maximum a-posteriori estimation into an energy minimization problem. Minimizing this energy function is equivalent to maximizing the total a posteriori probability of all voxel assignments.

This method takes less time to converge, which is about a factor 3 times faster than its previous methods. In comparison, the probabilistic derivation and formulation of the energy on the volumetric instead of the image domain provides faster convergence and better robustness. But this approach is susceptible to noise and shading effects, since single observations are taken into account for deciding whether a voxel should be categorize forward and backward voxel. This could be overcome by considering the averaging effect of integrating data from all views or by allowing user interactions

The same authors at [4] extended their work from this basic framework by adding user interaction to distinguish between the foreground and background. The user scribes with 2 different colors in order to distinguish between foreground and background. This avoids the shading effect existed in the previous work. It also uses the recent technology of parallel processing with GPU(A graphics processing unit also occasionally called visual processing unit or VPU) is a specialized circuit designed to rapidly manipulate and alter memory in such a way so as to accelerate the building of images in a frame buffer intended for output to a display. It consists of 100s of cores to execute. With this technology the time of processing gets reduced considerably up to a few seconds.

Jean Sebastian and Franco Edmond Boyer at [5] describe a Bayesian framework for multi-view silhouette. This considers pixel based information rather than considering voxel. This approach starts by having two assumptions. One is a set of images, obtained from fully calibrated cameras. Second is a set of background images of the scene, free from any object of interest, have previously been observed for each camera. Thus, we must first model the impact of joint probability on the observations. Modeling the relationships between the variables involved requires computing the joint probability of these variables. The decomposition is based on the statistical dependencies.

This very basic framework is extended and modified by Won woo et.al. in [6]. They specify a method for extracting consistent foreground regions when multiple views of a single scene are available. This automatically identifies such regions in images under the assumption that, in each image, background and foreground regions present with different color properties. Second assumption is that the object to be extracted includes entirely in each images.

In this method, one object is viewed from different angles and this different angle pictures are used in the process of silhouette extraction. In the first iteration a probable area is selected with the help of a histogram. The Bayesian frame work is utilized in this point of silhouette initialization. This is also an iterative mechanism. On iterations the probable area first selected gets refined and shrinks in to the shape of the object. This method utilizes graph cut method for iterative optimization. The graph cut approach finds new silhouette labels from which new background models are inferred before the next iteration. To terminate the iterative optimization, observe the number of pixels whose states changed from 'Unknown' to 'Background' and stop the process when no further pixels are newly identified as being in the background.

This is a fully automatic procedure which does not need any background information or user interventions. This method has some limitations; it fails when the background and the foreground have less color differentiation and when the object of interest does not appear entirely in each image. The first disadvantage can be overcome by integrating any of the finest edge detection mechanisms in this method.
The following table summarizes 4 different approaches of Bayesian frame work

Author	No. of images & Method	Time taken & Remark
K. kolev et al. 2006	20 Noisy, voxel based Bayesian	20-30 mints & automatic
K. k et al. 2011	27 Noisy, voxel based Bayesian	5 sec(GPU)& user interactive
E. Boyer et al. 2006	8, pixel based Bayesian	13 sec& user interactive
W. W .lee et al. 2011	6 consistent, Pixel based Bayesian	several mints & automatic

IV. CONCLUSION AND FUTURE WORK

Segmentation is used as a preliminary operation in most of the image processing applications. This covers a wide range of practical applications and computer vision applications. Real time applications of image processing-like satellite and medical image processing, especially needs the clear and clever extraction of objects from images. Multi-view image segmentation is best implemented with Bayesian network because it involves the use of inferred probability rather than observed one. The information from one image to the next is transmitted in each view of segmentation through Bayesian network and thus uses less time during successive iterations compared to multiple single view segmentations. Multi-view segmentation gives better results when graph cut is applied. The result is even finer and faster if provide with normalization graph cut. In normalization procedure a pixel changes its status to foreground if and only if its half or more than half neighbors are having foreground status. This method reduces the iterations needed to find the falsely identified pixels and hence false alarm rate (FAR).

B. Limitations and future work

This method fails to provide better result when image include complex background. Likewise if the image contains 2 or more objects, user interaction may needed to locate exact object. The future work can be extended to create a better statistical model for foreground having more than 2 regions. It can also be extended to multi-view video sequences.

V. REFERENCES

[1] S Sapna Varshney, Navin Rajpa and Ravindar Purwar "comparative Study of Image Segmentation Techniques and Object Matching using Segmentation" International Conference on Methods and Models in Computer Science ,2009.

[2] Dr .G .Padmavathi, M.Muthukumar, Suresh kumar Thakur "Implementation and Comparison of Different segmentation Algorithms used for Images Based on Nonlinear Objective Assessments" 3rd International Conference on Advanced Computer Theory and Engineering(ICACTE),2010.

[3] K. Kolev, T. Brox, and D. Cremers " Robust variational segmentation of 3D Objects from multiple views". In K. Franke et al., editor, Pattern Recognition Conference at Berlin (Proc. DAGM), volume 4174 of LNCS, pages 688697, Berlin, Germany, September 2006.

[4] K. Kolev, T. Brox, and D. Cremers. "Fast Joint Estimation of Silhouettes and Dense 3D Geometry from Multiple Images " Ieee transactionson pattern analysis and machine intelligence ,Digital Object Indentifier .1109/TPAMI.2011.150 0162, june 2011.

[5] J.-S. Franco and E. Boyer. "Fusion of multi-view silhouette cues using a space occupancy grid". In Proc. International Conference on Computer Vision, Beijin, China, 2005.

[6] Wonwoo Lee, Woontack Woo "silhoutte segmentation in multiple views " Ieee transactions on pattern analysis and machine intelligence, VOL. 33, NO. 7, 2011.The template will number citations consecutively within brackets

[7] Morris, O.; M. Lee “A unified method for segmentation and edge detection using graph theory “Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '86. Volume: 1986

[8] Juanying Xie; Shuai Jiang; “A Simple and Fast Algorithm for Global K-means Clustering” IEEE Education Technology and Computer Science (ETCS), 2010 Second International Workshop on Volume: 2 , 2010

[9] Guang Yang; Kexiong Chen; Maiyu Zhou; Zhonglin Xu; Yongtian Chen; Study on Statistics Iterative Thresholding Segmentation Based on Aviation Image , 2007. Eighth IEEE ACIS International Conference : 2007

[10] Hong Lan; Ling Chen; Wei Hu; “An apIproach on liver medical image

segmentation based on quad tree “Multimedia Technology (ICMT), 2011 IEEE International Conference 2011 .

[11] Xiaomu Song; Guoliang Fan; “A study of supervised, semi-supervised and unsupervised multiscale Bayesian image segmentation” IEEE international conference 2002 Page(s): II-371 -II-374 vol.2

[12] Morris, R.; Descombes, X.; Zerubia, J.; “Fully Bayesian image segmentation-an engineering perspective “Image Processing, Proceedings., IEEE International Conference on Pattern Recognition 1997

[13] Yi Zhou; Wei Zhang; Xiaoou Tang; Harry Shum; “A Bayesian mixture model for multi-view face alignment “Computer Vision and Pattern cognition, 2005. CVPR 2005. IEEE Computer Society conference 2005

[14] "Dino, Data Set," Multi-View Stereo Evaluation Web page, http://vision. middlebury.edu/mview /, 2010.

[15] Z. Wu and R. Leahy, "An Optimal Graph Theoretic Approach to Data Clustering: Theory and Its Application to Image Segmentation" IEEE Trans. Pattern Analysis
and Machine Intelligence, vol. 15, no. 11, pp. 1,101-1,113, Nov. 1993.

[16] R. C. Gonzalez and R. E. Woods. Digital Image Processing. Addison Wesley, 2nd edition, 1992.

[17] "Temple, Data Set," Multi-View Stereo Evaluation Web Page, http://vision.middlebury.edu/mview/, 2010.

[18] Y. Boykov and M.-P. Jolly, "Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images," Proc. Intl Conf. Computer Vision, vol. 1, pp. 105-112, 2001.

[19] C. Rother, V. Kolmogorov, and A. Blake, "Grab Cut-Interactive Foreground Extraction Using Iterated Graph Cuts," Proc. ACM SIGGRAPH, vol. 24, no. 3, pp. 309-314,
2004

[20] Jianbo Shi and Jitendra Malik "Normalized Cuts and Image Segmentation" IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,
VOL. 22, NO. 8, AUGUST 2000

[21] "Kung-Fu Girl, Data Set," http://www.mpi-inf.mpg.de/departments/irg3/kungfu/,
2010.