International Journal Of Scientific & Engineering Research, Volume 4, Issue 5, , Èà-2013
ISSN 2229-5518

Efficient Foreground Object Detection using Background Subtraction

K.Suganya Devi, N.Malmurugan, G.Selvasanthi

Abstract- Background subtraction is a computational vision process of extracting foreground objects in a video sequence. There are many algorithms that had been developed to perform background subtraction. In this paper, a cone shaped hierarchical background model (CSHBM) is proposed to extract the moving objects from the video sequence by incorporating more features from different previous background subtraction methods. A mean-shift algorithm is first applied to segment the background images into set of regions. Then CSHBM model is constructed which consists of two different models namely, region model and pixel model. Gaussian mixture model (GMM) is used to build the region models. This GMM is also used to fix a threshold values for cone shaped illumination model (CSIM). The CSIM is used to build the pixel models. Because of background image segmentation, different parameters can be set to different regions. Experimental results are presented which show that compared to existing techniques, the proposed method provides more robust segmentation in the presence of illumination changes.

Index Terms- Background Subtraction, Foreground Extraction, Region Segmentation, Pixel Model, Gaussian Mixture Model, Color-based Background Model, Cone Shape Illumination Model.

———————————— ◆ ————————————

1 INTRODUCTION

In computer applications, it is necessary to extract the foreground objects in a video sequence [1], [2]. Background subtraction is one of the methods used for foreground object extraction. In this technique a background model is constructed to detect the moving objects. This background model represents the scene with no moving objects. Each frame of the video is compared with background model. If the frames contain pixels that are not in the background model, then that pixels are considered as foreground. There are many subtraction methods that had been developed such as Mixture of Gaussians [3], [4], Kernel density estimation (KDE) [5], [6] and the co-occurrence of image variations [7].

While developing a good background subtraction algorithm, the researchers faced many challenges. Major challenges are as follows: Illumination variation due to environmental effect, Reflecting surfaces, Shadows, Flickering of light sources and monitors and Noise. These challenges must be accounted while developing a background subtraction algorithm.

In this paper, a Cone Shape Hierarchical Background Model (CSHBM) is proposed. Initially, a mean-shift algorithm [8] is used to segment the background images in the training set into set of regions. Then the region model and pixel model is constructed based on the Gaussian Mixture Model (GMM) and Cone Shape Illumination Model (CSIM) respectively. The region model acts as a first level detector to decide which region contains the foreground objects. Then the pixel model is used to locate the position of the foreground objects. The Gaussian Mixture Method is used to fix a threshold value for CSIM method.

The proposed method is quite different from the existing background subtraction methods. In the method proposed by Piccardi and Jan [9], the mean-shift is used at the pixel level, while in our method it is used to pre-segment the background images into set of regions. In the method proposed by Wu et al [10], the fixed regions are used while in our method dynamic regions are used to represent the dynamic scene effectively.

2 RELATED CONTRIBUTIONS

Much work had been done towards obtaining the best possible background model which works in real time. Stauffer and Grimson [3] describe a method which adaptively models each pixel as a Mixture of Gaussians. This method could deal with slow changes in illumination, repeated motion from the background cluster and long term scene changes. Zivkovic [12] present how the number of components can be selected online and improve the algorithm presented in [3]. Kaewtrakuplong et al. [13] present a method which improves the adaptive background mixture model by re-investigating the update equations. Javed, Safique and Shah [14] present a method which uses both color and gradient feature vectors along with GMM method to make the process of background subtraction more robust towards sudden illumination changes. Hanzi and Suter [15] evaluate the performance of the GMM method [3] and propose certain modification in color feature vector

- *K.Suganya Devi is currently working as Asst. Prof. in Computer Science and Engineering, University college of Engineering, Panruti. Email: ssuganya.ucep@gmail.com.*
- *N.Malmurugan is a Director of Mahindra Groups of Institution, Tiruchengodu. Email: n_malmurugan@yahoo.com.*
- *G.Selvasanthi is currently doing master degree in Computer Science and Engineering, University college of Engineering, Panruti. Email: selvasanthi@gmail.com*

International Journal Of Scientific & Engineering Research, Volume 4, Issue 5, May-2013
ISSN 2229-5518

to improve the performance of the algorithm. [16] and [17] do an error analysis and experimental validation. Jwu Shang [18] proposes a new 3D Cone Shape Illumination Model (CSIM) that deals with shadow and highlight removal. This algorithm deals more specifically with the indoor environment and claims to do well in the varying illumination and sudden light changing conditions. [19] and [20] also concentrate on detecting the foreground object in varying lighting and shadow removal. Bascle et al [21] present a new approach for automatic image color correction based on statistical learning. Yokoyama et al. [22] describe an approach for the detection and tracking of moving objects using lines computed by a gradient based optical flow and an edge detector.

3 OVERVIEW OF THE PROPOSED METHOD

The proposed method involves two levels of processing. In the first level, training set background images are segmented into regions by mean-shift algorithm [8]. In second level, pixel analysis is done to detect the foreground objects.

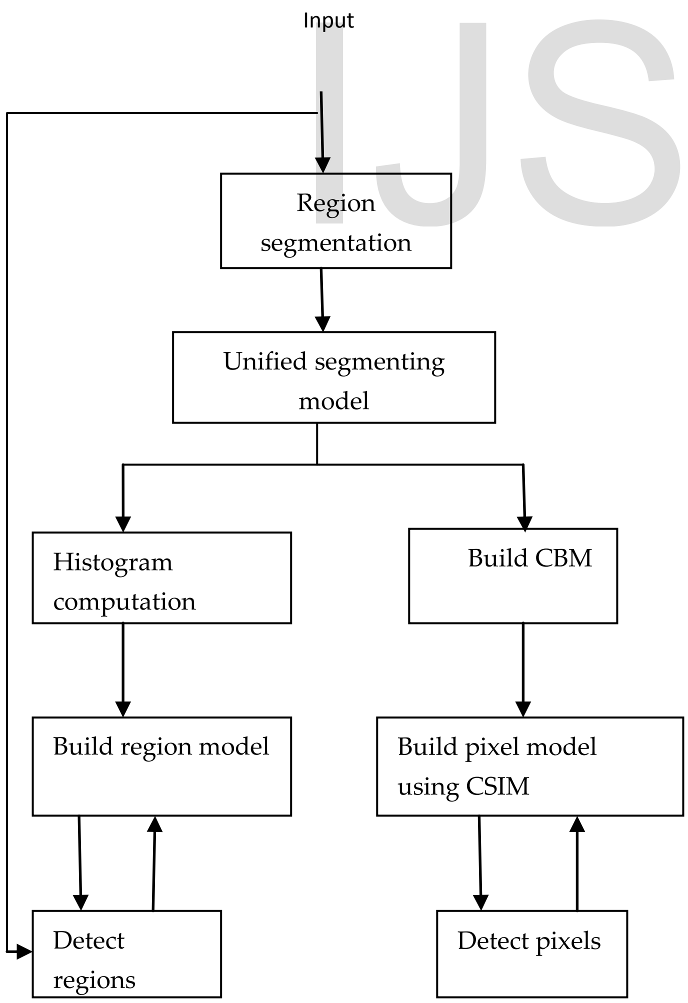

Fig. 1. Block diagram of the proposed method

Fig. 1. shows the block diagram of the proposed method. First the frames of the video are segmented into set of regions by mean-shift algorithm. Then regions are merged together to form a uniform segmenting model for a scene. Then the region models are constructed using gray histogram as features and pixel models are created based on color and gradient features. Gaussian mixture model is used to build the region model. For foreground object detection, first an input frame is segmented by unified segmenting model. Next region model is used to detect which regions contain foreground objects. If the region contains foreground objects, then the pixel models are used to locate the position of the foreground objects.

4 REGION MODELS

4.1 Region Segmentation

In the proposed method, initially the background images in the training set are segmented into set of regions by mean-shift algorithm [8]. Even if the camera is static, the segmented results may differ among frames due to environmental and light changing. The moving foreground objects also lead to different segments. When a new frame arrives, we need to segment it into regions and detect the regions for foreground objects. So we need a unified segmenting model to segment a new frame. Therefore the segmented results of the background images in the training set are merged together to build the unified segmenting model.

For merging the regions, first we compute the intersecting area between each region in the first background image and regions in other background images. If the area is larger than a threshold value, then the two regions are integrated as one unified regions. The weights of pixels are increased in the intersecting area. In this way a unified segmenting model is built.

4.2 Region Models

In order to build the region model, the histogram of each region in each background image can be calculated using normalized weight of the pixel. Weight of the pixel means that number of occurrence in one of the region over all the background images in the training set. The following equation is used to compute the histogram of the i^{th} region,

$$h^i(j) = K \sum w^i(x,y), p(x,y) \epsilon bin(j) \quad (1)$$

where $p(x,y)$ is the pixel value of point (x,y) and $bin(j)$ is the range of pixel values of j^{th} bin. K represents the summation of all bin values in the histogram to one.

International Journal Of Scientific & Engineering Research, Volume 4, Issue 5, May-2013
ISSN 2229-5518

Gaussian mixture model for each region is built in order to construct the region model. In the proposed methods number of distributions is different for each region. AP cluster algorithm [23] is used to compute the number of distributions of each region. Bhattacharyya distance can be calculated in order to measure the histogram similarity. This can be represented by

$$d_{nm}^{k} = \sum_{i=1}^{L} \sqrt{h_i^n h_i^m} \quad (2)$$

The mean of the Bhattacharyya distances of all histograms of the same region is used to fix a number of distributions. If the number of components of the i^{th} region is C^i and entries of the k^{th} component is e_k^i , then the weight of this component can be denoted by

$$w_k^i = e_k^i \,/\, M \quad (3)$$

Bhattacharyya distances among all entries are used to fit the normal distribution. After fitting the normal distribution for the components of all regions, we obtain the region model.

4.3 Foreground Detection

For detecting foreground objects, segment an input frame according to the unified segmenting model. Then each region is detected if it contains any foreground objects using a corresponding region model. The probability of the region that belongs to background is measured by

$$P_{reg}^{k} = \sum_{i=1}^{C^k} W_i^k \sqrt{F(d_i^k, G_i^k)[1 - F(d_i^k, G_i^k)]} \quad (4)$$

where $F(d_i^k, G_i^k)$ is a normal cumulative distribution. We can set a uniform threshold value for all regions to decide whether the region contains foreground or not.

$$bg^k = \begin{cases} 1, P_{reg}^k \geq T_r \\ 0, P_{reg}^k \leq T_r \end{cases} \quad (5)$$

If the value of bg^k is 1, then the region is background and if it is 0, then the region contains foreground. Then constant uniform threshold value is 0.22.

5 PIXEL MODEL

5.1 Color- based Background Model (CBM)

For constructing CBM model, Gaussian mixture model described in [3] is used. In CBM method, each pixel x is defined as a 3-dimensional vector (R, G, B) at time t. This is modeled by N Gaussian distributions. To improve the flexibility of the CBM, Long term Color-based Background Model (LTCBM) is defined with extra new N Gaussian distributions.

$$B = \underset{b}{argmin} \sum_{j=1}^{b} w_j > B_0 \quad (6)$$

Where $b \leq N$ and B_0 is the threshold value. To determine the background, the first b Gaussian distributions are used which are defined in the ECBM and the remainders (2N-b) of the Gaussian distributions are used for dealing with background changes which are defined in the CCBM. These two ECBM and CCBM are combined to define a LTCBM.

The background changes during short period B_1 are recorded using Short Term Color-based Background Model (STCBM). A pixel value set $PV = \{P_1, P_2, \dots \dots P_k, \dots . P_{B_1}\}$ collected during a period B_1, the corresponding Gaussian distribution set $CG = \{g_1, g_2, \dots \dots g_k, \dots . g_{B_1}\}$ is calculated by comparing the pixel value set with LTCBM. The histogram of CG is calculated by

$$H_{CG}(Z) = \sum_k \delta(z - g_k) \,/\, B_1 \quad (7)$$

A transfer flag set F_{CG} is defined for adjusting the bin order in $H_{CG}(Z)$ as follows,

$$F_{CG} = \left\{F_j, j = 1, \dots, B, F_j \in \{-2N + 1, \dots, 0, 1, \dots, 2N + 1\}\right\} \quad (8)$$

5.2 Cone Shape Illumination Model (CSIM)

In the RGB color space, a Gaussian distribution in the LTCBM becomes an ellipsoid whose center is the mean of the Gaussian component and the length of each principle axis equals 2.5 standard deviations of the Gaussian component. A new pixel $I(R, G, B)$ is considered to belong to background if it is located inside the ellipsoid. The chromaticities of the pixels located outside the ellipsoid but inside the cone resemble the chromaticity of the background. The brightness difference is then applied to classify the pixel as either highlight or shadow.

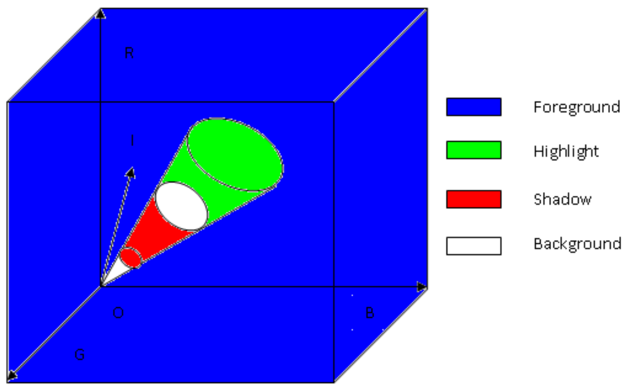

Fig. 2 CSIM in the RGB color space

International Journal Of Scientific & Engineering Research, Volume 4, Issue 5, May-2013
ISSN 2229-5518

The threshold values τ_{low} and τ_{high} are applied to classify the pixel value as shadow or highlight and that can be selected based on the standard deviation of the corresponding Gaussian distribution in CBM. Fig. 2 shows the proposed Cone Shape Illumination Model in the RGB space.

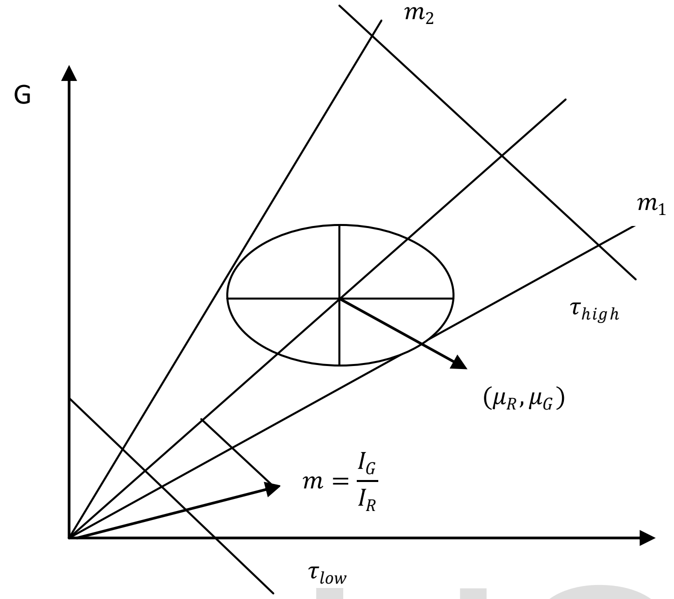

Fig. 3 2D projection of the 3D CSIM from RGB space to RG space

The ellipse center is (μ_R, μ_G) and elliptical equation is denoted by

$$(R-\mu_R)^2/a^2 + (G-\mu_G)^2/b^2 = 1 \qquad (9)$$

Where $a = 2.5 * \sigma_R$ and $b = 2.5 * \sigma_G$.

$$m_{1,2} = \frac{-(2\mu_R\mu_G) \pm \sqrt{(a^2-{\mu_R}^2)^2 - 4(2\mu_R\mu_G)(b^2-{\mu_G}^2)^2}}{2(a^2-{\mu_R}^2)^2} \qquad (10)$$

A matching result set is given by $F_b = \{f_{bi}, i = 1,2,3\}$, where f_{bi} is the matching result of a specific 2D space. A pixel vector $I = [I_R, I_G, I_B]$ is projected onto the 2D spaces of $R-G, G-B$ and $B-R$. The pixel matching result is set to 1 when the slope of the projected pixel vector is between m_1 and m_2.

If the background mean vector is $E = [\mu_R, \mu_G, \mu_B]$, the brightness distortion $\propto_b$ can be calculated by

$$\propto_b = |I| \cos\left(\left|\tan^{-1}\left(\frac{I_G}{\sqrt{{I_R}^2+{I_B}^2}}\right) - \tan^{-1}\left(\frac{\mu_G}{\sqrt{{\mu_R}^2+{\mu_B}^2}}\right)\right|\right) / |E| \qquad (11)$$

The image pixel is classified as highlight, shadow or foreground using the matching result set F_b, the brightness distortion $\propto_b$

$$C(I) = \begin{cases} Shadow\colon \sum F_b = 3 \text{ and } \tau_{low} < \alpha_b < 1, else \\ Highlight\colon \sum F_b = 3 \text{ and } 1 < \alpha_b < \tau_{high}, else \\ Foreground\colon otherwise \end{cases} \qquad (12)$$

τ_{low} and τ_{high} can be chosen using N_G standard deviation of the corresponding Gaussian distribution in CBM and are described as

$$\tau_{high} = 1 + \sqrt{(N_G \cdot \sigma_R)^2 + (N_G \cdot \sigma_G)^2 + (N_G \cdot \sigma_B)^2} \cdot \cos\theta_r \cdot L_\mu \qquad (13)$$

$$\tau_{low} = 1 - \sqrt{(N_G \cdot \sigma_R)^2 + (N_G \cdot \sigma_G)^2 + (N_G \cdot \sigma_B)^2} \cdot \cos\theta_r \cdot L_\mu$$

where

$$\theta_r = |\theta_E - \theta_S| = \left|\tan^{-1}\left(\frac{\mu_G}{\sqrt{{\mu_R}^2+{\mu_B}^2}}\right) - \tan^{-1}\left(\frac{\sigma_G}{\sqrt{{\sigma_R}^2+{\sigma_B}^2}}\right)\right| \qquad (14)$$

$$L_\mu = 1 / \sqrt{{\mu_R}^2+{\mu_G}^2+{\mu_B}^2} \qquad (15)$$

$C(I)$ is defined as the result of color-based background subtraction using CBM. The foreground pixels labeled in $C(I)$ are further classified as shadow, highlight and foreground by using CSIM. $C'(I)$ can then be obtained from $C(I)$ after transferring the foreground pixels which have been labeled as shadow and highlight in $C(I)$ into the background pixel.

5 RESULTS AND DISCUSSION

The Cone Shaped Hierarchical Background Model (CSHBM) is implemented on a personal computer with a 1.73 GHz Intel Pentium dual-core processor and 2GB RAM, using MATLAB. This algorithm is implemented based on two stages namely training stage and detection stage.

To analyze the complexity of the proposed method, we estimate the time cost in each phase of two stages. The training stage consists of four phases like frame segmentation, merging regions, building the region model and pixel model. Region segmentation is a time consuming phase using mean shift to segment a large frame. So all frames are resized to 160 × 120. Thus the average time for frame segmentation ranges from 0.042 to 0.056s. The merging regions phase is based on these resized frames to save time. In the training stage, this phase is executed once. This phase takes 0.5s time. The region model is built using Gaussian mixture method and AP cluster algorithm is used to fix the number of components. These two algorithms are very

International Journal Of Scientific & Engineering Research, Volume 4, Issue 5, May-2013
ISSN 2229-5518

fast. So this phase takes short time that ranges from 1.8 to 2.5s. The building pixel models also take short time that ranges from 2.7 to 3.2s.

The detection stage contains two phases namely region model detection and pixel model detection. In our method first detects the regions that contains background and neglect the regions without background. A large number of regions cause high processing time for detection of region model. Detection of pixel models is only in the detected regions. Thus the processing time for this phase is short. Fig. 4 shows the detection of foreground object in a test video.

Fig. 4. Results of foreground detection from test video. From left to right, original frame and the results of our method respectively.

6 CONCLUSION

In this work, we implement a Cone Shaped Hierarchical Background Model (CSHBM). Hence we used two levels of models for detecting foreground objects. First the region models are used to decide which region contains foreground. Then the pixel models are used to locate the foreground objects in the certain region resulted by the region model. Thus processing time is reduced for foreground object detection.

In this work, we use a Cone-Shape Illumination Model (CSIM). Thus, the shadows and highlights are removed. So, we get an accurate foreground objects.

REFERENCES

[1] Y. Chen, B.Wu, H. Huang, and C. Fan, "A real-time vision system for nighttime vehicle detection and traffic surveillance," IEEE Trans. Ind. Electron., vol. 58, no. 5, pp. 2030–2044, May 2011.

[2] B. Armstrong and S. PuthanVeettil, "Soft synchronization: Synchronization for network-connected machine vision systems," IEEE Trans. Ind. Informat., vol. 3, no. 4, pp. 263–274, Nov. 2007.

[3] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., Jun. 1999, pp. 246–252.

[4] D. S. Lee, "Effective Gaussian mixture learning for video background subtraction," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 5, pp. 827–832, May 2005.

[5] A. Elgammal, D. Harwood, and L. S. Davis, "Non-parametric model for background subtraction," in Proc. Euro. Conf. Comput. Vis., Jun. 2000, pp. 751–767.

[6] A. Mittal and N. Paragios, "Motion-based background subtraction using adaptive kernel density estimation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2004, pp. 302–309.

[7] M. Seki, T. Wada, H. Fujiwara, and K. Sumi, "Background subtraction based on cooccurrence of image variations," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2003, vol. 2, pp. 65–72.

[8] D. Comaniciu and P. Meer, "Robust analysis of feature spaces: Color image segmentation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 1997, pp. 750–755.

[9] M. Piccardi and T. Jan, "Mean-shift background image modelling," in Proc. Int. Conf. Image Process., 2004, vol. 5, pp. 3399–3402.

[10] P. Varcheie, M. S. Lavoie, and G. A. Bilodeau, "An efficient region- based background subtraction technique," in Proc. Canadian Conf. Comput. Robot Vis., 2008, pp. 71–78.

[11] W. Hu, H. Gong, S. Zhu, and Y. Wang, "An integrated background model for video surveillance based on primal sketch and 3D scene geometry," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2008, pp.1–8.

[12] Z.Zivkovic,"Improved adaptive Gaussian mixture model for background subtraction", IEEE Int. Conf. Comput.Vis. Pattern Recognit., 2004, pp.28-31.

[13] P.Kaewtrakuplong and R.Bowden, "An improved adaptive background mixture model for eal-time tracking and shadow detection", In Proc. Europ. Workshop on Advanced Video based surveillance systems, comp.vis. and distributed processing, 2001.

[14] O.Javed, K.Safique and M.Shah,"A hierarchical approach to robustbackground subtraction using color and gradient information", 2002,pp.22-27.

[15] Hanzi Wang and D.Suter,"A re-evaluation of mixture of Gaussian background modeling", IEEE Int.Conf.onAcou.Speech and Signal Process.,2005,pp.1017-1020.

[16] M. Piccardi, "Background subtraction techniques: A review," in Proc. IEEE Int. Conf. Syst., Man Cybern., 2004, vol. 4, pp. 3099–3104.

[17] Xiang Gao, T.E.Boult, V.Ramesh and F.Coetzee, "Error analysis of background adaption",IEEE Conf. on Comp.Vis. and Pattern Recognit., 2000.

[18] Jwu-Sheng Hu,Tzung-Min Su and Shr-Chi Jeng,"Robust Background Subtraction with shadow and highlight removal for indoor surveillance", IEE/RSJ,Int.Conf. onIntelligent Robots and Systems,2006,pp.4545-4550.

[19] Jianguang Lou, Hao Yang, Weiming Hu and TieniuTan,"An illumination invariant change detection algorithm",Asian Conf. on Comput.Vis.2000.'

[20] T.Horprasert,D.Heirwood and Larry S.Davis,"A robust background subtraction and shadow detection", Asian Conf. on Comput.Vis.,2000.

[21] B.Bascle,O.Bernier and V.Lemaive,"A statistical approach for learning invariant application to image color correction and learning invariants to illumination", Int.Journal of Imaging Syst. And Tech.,2007

[22] Yokoyama,"A contour-based moving object detection and trackling",2005.

[23] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," Science, vol. 315, pp. 972-976, 2007.