

# Clustering Image Sequence Based on Potts Model of the Motion Field

Keisuke Kinoshita \* Iris Fermin †‡

## Abstract

The computation of visual motion at multiple locations is important for tasks such as stereo motion, depth estimation, structure from motion, and control of locomotion. From an image sequence we can classify objects according to their motion; grouping static objects and moving objects, for instance. Such classification of image frame regions can be used for tasks such as surveillance systems and attention control for humanoid robots. In this paper we present an approach for grouping image objects by motion decomposition, based on the Potts model [1] and Monte Carlo simulation of the spatial-temporal information. The temperature changes in the Potts model allow clustering the spins (pixels), thus at low temperature the spins that belong to the same cluster are aligned.

*Key words:* motion segmentation, region clustering, correlation, Potts model, Monte Carlo simulation

## Potts モデルによる動画像のクラスタリング

木下敬介 \* イリス・フェルミン †‡

### あらまし

視覚による動きの検出は、ステレオ視や奥行き推定、3次元復元、自立歩行などのタスクに必須である。人間は、画像列の中から対象物体を、その動きに基づいて分類（たとえば、静止物体と運動している物体）することができる。このような機能を実現できたなら、監視システムや人型ロボットの注意生成機構に応用可能である。本研究では、Potts モデルとモンテカルロ・シミュレーションに基づいた動きの分類手法を提案する。Potts モデルは、磁性体のスピンモデルの拡張であり、温度をパラメータとして持つ。この温度パラメータを変化させると、スピンの結合具合が変化する性質を利用して、画像列の動きを抽出する。

キーワード: 動き分割, 領域分割, 相関, Potts モデル, モンテカルロ・シミュレーション

## 1 Introduction

One of the most challenging issues for computer vision is to develop visual mechanisms that facilitate the interaction with dynamic environments. Object motions that repeat are common in both nature and man-made environments. Moving objects over background surfaces that may themselves be mobile, require the computation of relative motion. Thus, use of motion as a cue can be a meaningful factor for object recognition and control in mobile robot applications.

Clustering objects given their motions has similarities with image segmentation, which is the division of an image frame into different regions each

having certain properties. The level to which these divisions of the image are carried depends on the problem being solved. This process is similar to the human recognition process [2, 3]. The extraction of the motion information is also categorized under three main approaches: trajectory-based features, optical flow and region-based features [4, 5, 6].

Since the problem of motion representation is closely related to image segmentation and efficient motion estimation solutions should be able of addressing both components. This remark has led to the proliferating of algorithms with iterate between optical flow estimation and segmentation [7, 8]. Such algorithms can be seen as variations of the expectation-maximization model or Markov Random Fields, which is very popular approach for modeling spatial interactions on lattice systems. An important characteristics of MRF modeling is that global patterns are formed via stochastic propagation of local interactions [9, 10], however, difficulties

\*ATR Human Information Science Research Labs.

ATR 人間情報科学研究所 E-mail: kino@atr.co.jp

†School Eng. & Applied Science, Aston University, UK  
アストン大学 英国 E-mail: i.fermin@aston.ac.uk

‡This work was developed during her visit to ATR, Dec. 2002

arise due to the dependence structure in the models and approximations are required to make the algorithm manageable.

Cremers et.al. [11] presented a variational method for the segmentation of piecewise of the flow fields, in which minimization of the motion energy and internal shape energy are computed. This allows that objects that are not discernible by their appearance can be segmented according to their motions. The drawback of this approach is that the model of the shape has to be assumed and the complexity increase as the model of the shape become more complex. Bayesian approaches are well known since their robustness to outliers but the estimation of the priors is computational expensive [9, 12, 13]. Approaches based on level sets that exploit the ability to handle variations in the topology of the segmentation and its numerical stability have also been proposed [14, 15].

## 1.1 Visual Motion Perception

Motion perception is fundamental for our understanding of and our communication with the environment. Visual motion can be a powerful cue that enables us to classify different forms of movement such as rigid versus non-rigid, as well as recognizing classes of objects on the basis of their motion characteristics. The ability of recognize classes of objects on basis of their motion characteristics is an essential requirement to build robust vision systems.

There are both psychological and physiological evidences that the visual system processes moving images using a bank of filters tuned to specific spatial and temporal frequencies [16]. It is well known that the points on the surface of an object, moving relative to a camera, generate trajectories in space-time. The projection of these trajectories onto the imaging surface produces a two dimensional path, the time derivatives of which are the two dimensional velocities.

Two main approaches for the computation of image motion have been studied extensively: optical flow [17, 18], and extraction of three dimensional scene structure from image velocities [5, 19].

The difficulty in measuring image velocity is that image intensity depends on several independent aspects of the image formation process. Horn [17] pointed out that the optical flow cannot be computed at a point in the image independently of neighboring points without introducing additional constraints, because the velocity field at each image point in the image plane has two components, and the brightness change has only one.

We propose a method based on the Potts model using correlation and flow information to cluster dif-

ferent moving objects. The spatial-temporal and brightness information is used to cluster the data. In the spin model every point is assigned a spin value that represents its state. The state of the spin varies in order with the spin-spin interaction.

Potts, or spin models, are classical examples of many body systems where local two body interactions increase the whole group behavior. The group behavior is determined by the fact that many of the spins are found in the same state. This state similarity represents a similarity measure based on the spin-spin correlation function. In contrast with the distance measure used in many methods, spin-spin correlation is an inherently group-like property and is influenced by the state of the spins of the neighboring regions [20]. The strength of this kind of approach is that general image structures can be modeled using similar graph structures for representation, whether the interactions are between neighboring pixels or regions.

In sections 2 and 3, the Potts model and clustering method are explained in detail. In section 4 the implementation of the algorithm and computational results are introduced. Finally, in section 5 the conclusions are presented.

## 2 Potts Model and Swendsen and Wang Algorithm

The Ising model is the simplest model for ferromagnetic material where the number of states ( $q$ ) that a spin can have are two (2). The model is originated from physical studies of magnetics [21]. One of the interesting feature is that it exhibits a spontaneous magnetization at low temperatures. At the same time, the magnetization disappears at a higher temperature,  $T_c$ .  $T_c$  is called the critical temperature and the phases below and above this temperature are called ferromagnetic and paramagnetic, respectively. The Potts model is a generalization of the Ising Model to more than two states [22].

Given a sequence of points  $0, 1, \dots, n$  in a line. At each point, or site, there is a spin which at given moment is in one of the two positions that it can take: up or down. In a two-dimensional space, for example a square lattice  $L \times L$ , we put a spin  $s_i$  (this may take several values in the Potts model) at each lattice site  $i$ . The set of  $L^2$  spins consists of the state space  $X$ . The system has a total energy defined by

$$\mathcal{H} = - \sum_{\langle i,j \rangle} J_{ij} \delta_{s_i s_j} \quad s_i = 1, \dots, q \quad (1)$$

where  $J$  is the interaction between spins and only interactions between neighbors are considered. For

$J > 0$  the system is in the ferromagnetic phase, and if  $J < 0$  the phase is anti-ferromagnetic. Changes in temperature will generate changes in the state of the system. The changes of the temperature define a probability distribution of the states:

$$P(s) \propto \exp\left(-\frac{\mathcal{H}(s)}{zT}\right) \quad (2)$$

where  $z$  is a normalization constant.

The Swendsen-Wang (SW) is a multi-cluster algorithm for Ising and Potts model. SW starts with a spin configuration  $s$  and generates a percolation configuration based on the following method:

- Pick an arbitrary state  $s$
- Go through each neighboring connection in the lattice, create a bond between two neighboring sites  $i$  and  $j$  with probability  $1 - e^{-J/zT}$  but only if the spins are the same.
- Identify clusters as a set of sites connected by bonds, or isolated sites. After a cluster is found, each cluster is assigned a new Ising spin chosen with equal probability. The old spin values are discarded.
- One Monte Carlo step is finished, repeat.

The method is such that the transition leaves the equilibrium probability invariant, i.e., the system is ergodic. The SW algorithm generates many clusters and flip them at the same time, this makes this algorithm very efficient [21, 23, 24]. Algorithms using similar strategy (annealing) are well-know in the area of texture analysis [25, 26] or motion estimation [27], given that this method possesses desirable global optimization properties.

### 3 Potts Model for Motion Clustering

The  $q$ -state Potts model consists of a lattice  $s_i$  which can take  $q$  values and is governed by a Hamiltonian energy Eq. 1. Based on the model proposed by Blatt et.al.[1], we developed an algorithm to cluster the image pixels given their motion fields. Having  $N$  motion fields we want to cluster all pixels in  $M$  groups that should be determined. We assumed that the number of moving objects is unknown and a perfect clustering will be reached if  $N = M$ .

We assign a spin variable  $s_i$  to each  $\mathbf{x}_i$  that represents our velocity pattern, and introduce an interaction  $J_{ij}$  between pairs of spin, whose strength decreases as the inter-spin distance  $d_{ij}$  decreases.  $\delta_{s_i s_j}$  is the spin-spin correlation. Let an image sequence  $F$  of  $P \times Q$  pixels  $f(x, y, k)$  by frame, where

$x$  and  $y$  refer to spatial coordinates and  $k$  refers to the temporal coordinate.

We define the vector  $\mathbf{x}_i = (f_x, f_y, f_t)$  derived from the gradient constraint [17]

$$\begin{aligned} f_x &= \frac{1}{4}[f_{i+1,j,k} - f_{i,j,k} + f_{i+1,j+1,k} - f_{i,j+1,k} \\ &\quad + f_{i+1,j,k+1} - f_{i,j,k+1} + f_{i+1,j+1,k+1} - f_{i,j+1,k+1}] \\ f_y &= \frac{1}{4}[f_{i,j+1,k} - f_{i,j,k} + f_{i+1,j+1,k} - f_{i+1,j,k} \\ &\quad + f_{i+1,j,k+1} - f_{i,j,k+1} + f_{i+1,j+1,k+1} - f_{i+1,j,k+1}] \\ f_t &= \frac{1}{4}[f_{i,j,k+1} - f_{i,j,k} + f_{i+1,j,k+1} - f_{i+1,j,k} \\ &\quad + f_{i+1,j,k+1} - f_{i,j+1,k} + f_{i+1,j+1,k+1} - f_{i+1,j+1,k}] \end{aligned} \quad (3)$$

and

$$d_i = \frac{-f_t}{\sqrt{f_x^2 + f_y^2}} \quad (4)$$

where  $d_i$  indicates the component of the movement in the direction of the gradient  $(f_x, f_y)$ . The velocity  $(v_x, v_y)$  has to lie along a line perpendicular to the brightness gradient vector  $(f_x, f_y)$ . The distance of this line to the origin is  $d_i$  (see Fig. 1). Thus,  $d_i$  can be used for cluster points that move with similar velocity.

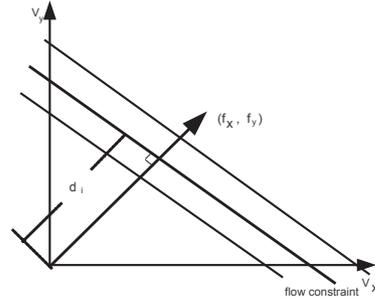


Figure 1: Image brightness equation constraint.

#### 3.1 Local Interaction and Thermodynamics Parameters

The local interaction is defined allowing only  $K$  neighbors and we assume a local length scale  $a$ . The interaction strength is [1]

$$J_{ij} = J_{ji} = \frac{1}{K} \exp\left(-\frac{|d_i - d_j|^2}{2a^2}\right) \quad (5)$$

Here  $a$  is set to the mean of the  $|d_i - d_j|$  in the neighbours,  $\widehat{K}$  is the average number of neighbors per site.

The spin-spin correlation affects the ordering properties of the system, and can be estimated using a Monte Carlo procedure. The SW algorithm is used to identify clusters of similar spins. Starting with an initial spin configuration, SW selects the points over the neighborhood given the following probabilities,

$$P_d^{i,j} = \exp\left(-\frac{J_{ij}}{T}\right) \quad \text{to delete a spin} \quad (6)$$

$$P_f^{i,j} = 1 - P_d^{i,j} \quad \text{to freeze a spin.} \quad (7)$$

Thus the points that belong to the same cluster are those which have a path of frozen bonds connecting them. A new configuration is generated, by assigning randomly a value  $s_i = 1, \dots, q$  to the spins of each cluster. This define one step of the Monte Carlo procedure and should be iterated until a thermodynamic equilibrium is reached. The relation of the SW clusters and the ordering properties of the Potts spins is given in the following relation

$$\bar{\delta}_{s_i, s_j} = \frac{(q-1)\bar{n}_{ij} + 1}{q} \quad (8)$$

where  $\bar{n}$  is the average over the temperature and

$$n_{ij} = \begin{cases} 1 & \text{if } s_i \text{ and } s_j \in \text{the same cluster} \\ 0 & \text{otherwise.} \end{cases}$$

The transition from ferromagnetic phase to paramagnetic phase [1] can be approximated as

$$T_c \approx \frac{e^{-\frac{1}{2}}}{4 \log(1 + \sqrt{q})} \quad (9)$$

and the susceptibility of the system is proportional to the magnetization

$$\chi = \frac{N}{T} (\overline{m^2} - \bar{m}^2) \quad (10)$$

where the magnetization,  $m$  is defined as

$$m = \frac{qN_{max}/N - 1}{q-1} \quad N_{max} = \max\{N_1, \dots, N_q\}, \quad (11)$$

and  $N_q$  is the number of spins with the value  $q$ . Hence the point where the susceptibility vanishes is an upper bound for the transition temperature from super-paramagnetic to the paramagnetic phase. Thus, we apply a Monte Carlo (MC) procedure to identify the temperature transition. The magnetization of the system tells us how well the spins are aligned. Perfect alignment corresponds to complete order in the system. When  $T$  increases, the order decreases, for  $T > T_c$  the order is lost.

## 4 Clustering Strategy

The procedure for clustering consists of two main steps, firstly the identification of the range of temperatures in which the cluster may be observed, secondly, identification of the clusters using the information given by the spin-spin correlation function at this temperature. However, before to proceed with the clustering a pre-filtering is applied to reduce the size of the data.

### 4.1 Image Space Reduction

Given an image sequence, let  $b = f(x, y, k)$  be the pixel brightness in the location  $x, y$  at the time  $k$ ,  $k = 0, \dots, \tau$ . If an object is moving with respect to the cameras plane, then a change between two consecutive image frames will be registered as a change of the brightness and position. Thus given a vector  $\mathbf{I}_k(x, y, b)$ , a correlation function is defined as follows.

$$c_{k,k+1}(x, y, b) = \frac{\mathbf{I}_k(x, y, b)\mathbf{I}_{k+1}(x, y, b)^\top}{\|\mathbf{I}_k(x, y, b)\| \|\mathbf{I}_{k+1}(x, y, b)\|} \quad (12)$$

such that

$$|c_{k,k+1}(x, y, b) - 1| > \epsilon \quad (13)$$

where  $\epsilon$  is a small number.

Eq. 13 allows us to reduce the searching space in the image region, and is known that a great percentage of an image frame corresponds to static background. Clearly, if  $c_{k,k+1}$  is zero then not changes were registered for the corresponding pixel. Using this constraint the process of clustering the moving pixels is accelerated.

### 4.2 Clustering Algorithm

The details of our algorithm which is based on the Blatt [1] clustering procedure, are as follows.

- 1 Select a pair of images, and apply a Gaussian filter to the image difference.
- 2 Applied the correlation filter Eq. 12 and select the respective image points to be clustered.
- 3 Using Eq. 3 estimate the points  $\mathbf{x}_i$  using  $3 \times 3$  neighbors.
- 4 Select an initial temperature interval using Eq. 9. Set  $T$  to  $T = 0.00$  until  $T = 2T_c$ , and divide it into intervals of  $Tsteps$ .  $Tsteps$  is set to 100 in our experiments.
- 5 Do until limit of iterations  $M$  has reached.  $M$  is set to 2000 in our experiments.
- 5.1 Assign  $q$ -states (uniform distributed) -Potts spin variable-  $s_i$  to each  $\mathbf{x}_i$  or cluster.
- 5.2 Compute the point interactions  $J_{ij}$  using Eq. 5.

- 5.3 Compute the probability of bonding by Eqs. 6 and  $\delta_{ij}$  by Eq. 8 for all sites.
- 5.4 Connect (bonding) or delete each neighbors for each site depending on the probability computed in step 5.3.
- 5.5 Draw a configuration identifying all the clusters. A cluster is defined by a group of sites which have bonds connecting each other.
- 5.6 Repeat from 5.
- 6 Calculate the susceptibility  $\chi$  using Eq. 10.
- 7 Identify the range of temperatures between the susceptibility density  $\chi T/N$  vanishes and  $\chi T/N$  is maximal, set the temperature for cluster  $T_{cluster}$  inside that interval.
- 8 Repeat steps 5 for the given the  $T_{cluster}$ .  $T_{cluster}$  is supposed to be the best clustering temperature.
- 9 Compute  $\bar{n}_{ij}$  and then estimate the spin-spin correlation  $\bar{\delta}_{s_i s_j}$  for all neighboring pairs  $\mathbf{x}_i$  and  $\mathbf{x}_j$  using Eq. 8.
- 10 The cluster is selected according with  $\bar{\delta}_{s_i s_j}$  and a threshold  $\theta$ . Here  $\theta$  is set to the value 0.5.

## 5 Computational Results

In this section we present two examples of clustering different motions: tennis sequence<sup>§</sup> and mouse sequence, respectively. The images used in the tennis sequence are shown in Fig. 2. These images are 240 by 352 pixels. We construct the points to be clustered using the correlation constraint and the estimated distance  $d_i$ . The total number of points to be clustered in this case are 2008. We conducted 2000 iterations of Monte Carlo steps through out the experiments. Figures4 shows the clusters at  $T = 0.00$ ,  $T = 0.049$ ,  $T = 0.091$ ,  $T = 0.134$ , respectively. For this simple case, the clustered regions are stable until the temperature is raised to high. The clustering process is susceptible to the selection of  $T$  as shown in Fig. 5.

In the second experiment, the mouse sequence is used as shown in Fig. 3. Here the hand, the purple mouse pad, the Pooh mouse pad, the cd-rom and the reel are moved at the same time. We rescale these images to 60 by 80 pixels and reduce the number of points to be clustered to 1601. Fig. 6 is the results for  $T = 0.00$  and  $T = 0.087$ ,  $T = 0.102$ , and  $T = 0.138$ , respectively. We can observe that at  $T = 0.102$  six clusters are found: reel, cd-rom, Pooh, hand, purple mouse, and some shadow effects. Fig. 7 shows the susceptibility of the clustering process.

<sup>§</sup>The images are taken from <http://sampl.eng.ohio-state.edu/sampl/database.htm>

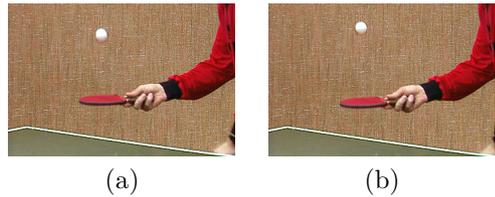


Figure 2: Two frames of the tennis sequence. (a) Image1. (b) Image2.



Figure 3: Mouse sequence 640x480 pixels: (a) Image1. (b) Image2.

The neighboring points were defined as eight points which surrounds the point. The parameter  $a_i$  that is a local scale parameter is selected as the mean of  $|d_i - d_j|$ , where  $j$  is the neighboring points of point  $i$ .  $\hat{K}$  which is a global scale was computed as the average number of neighboring points over all the points. The interval to select  $T$  was set between  $T_{min} = 0.0$  and a maximum of  $2T_c$ , where  $T_c$  is given by Eq.9. The threshold  $\theta$  for the selection of the cluster is set to 0.5. We notice that the selection of  $T_{cluster}$  as the average of the  $T$  between  $T_{vanish}$  at  $\chi T/N$  vanishes and  $T_{peak}$  at  $\chi T/N$  is maximal is not necessarily a good estimator. But the selection of  $T_{cluster}$  is inside that interval and around  $T_c$ . For the tennis sequence  $T_{cluster}$  is almost the same as  $(T_{peak} + T_{vanish})/2$ . But for the mouse sequence it is not a good estimator. We are studying the influence that flow data considered has on the selection of  $T_{cluster}$ .

Figure 8 shows a sequence of the results for the mouse sequence, but in this case the parameter  $a$  is estimated using the *median* values of  $|d_i - d_j|$ . Fig 9 shows the susceptibility density  $\chi T/N$  in which can be noticed that the phase transition is reached faster than that using the mean value. Finally, Fig. 10 is a synthetic color image data to explain the possibilities of the extension of the method to higher dimensional vector. In this case r,g,b intensities are used to compute  $d_i$ , and good segmentation results are shown in Fig. 12. Fig. 11 shows its susceptibility density  $\chi T/N$ . Thus, it is feasible to extend this method to work with more than one gradient constraint to deal with problems related to apparent motion, and occlusion without considering shape

models.

## 6 Conclusion and Discussion

We have presented an algorithm based on Potts model and Monte Carlo sampling for the clustering of several motions. Monte Carlo models are useful for obtaining numerical solutions to problems which are complicated to solve analytically. The error scale of MC models like  $1/\sqrt{(N)}$ , is independent of the number of dimensions. This makes this kind of method very attractive. Optical flow fields are inherently uncertain because of image noise, lighting changes, low contrast regions, the aperture problem and multiple motions in a single localized region. A probabilistic approach allows to represent these uncertainties in the computation. In our experiments, using a one-dimensional measure of the flow we were able to determine different objects in the sequence.

We found that changes in the temperature  $T$  lead to changes in the cluster distribution. For low temperatures, there are few clusters with high point density. And for high temperatures there are many clusters with low point density. The temperature parameter is a control scale parameter that defines the fitness of the cluster space. However, we need to find a robust way to select the  $T_{cluster}$ . Likewise, we would like to extend the method to deal with 2- and 3-dimensional measure of the flow.

## 7 Acknowledgements

One of the authors (IF) is grateful to Manfred Opper at Aston University for many helpful discussions, and would like to thank ATR Human Information Science Laboratories for their hospitality during part of this work. The research reported here was supported in part by a contract with the Telecommunications Advancement Organization of Japan entitled, 'Research on Human Communication'.

## References

- [1] M. Blatt, S. Wiseman, and E. Domany. Clustering data through an analogy to the potts model. In *Neural Information Processing Systems, NIPS*, volume 8, 1995.
- [2] M.I. Posner and M.E. Raichle. *The Images of Mind*. W.H. Freeman, 1994.
- [3] Z-L. Lu, C. Liu, and B. Doshier. Attention mechanisms for multi-location first- and second-order motion perception. *Vision Research*, 40:173–186, 2000.

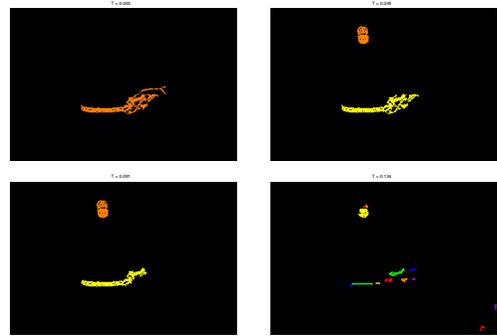


Figure 4: Results using the tennis sequence at different  $T$ .

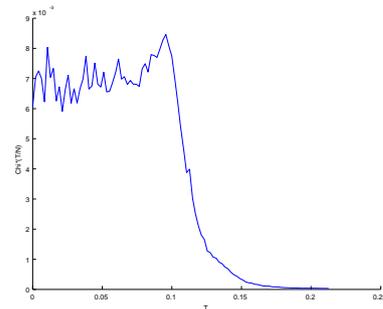


Figure 5: Tennis data: The susceptibility density of the clustering given different temperatures  $\chi^2 T/N$  vs.  $T$ , the parameter  $a$  computed as the *median*.

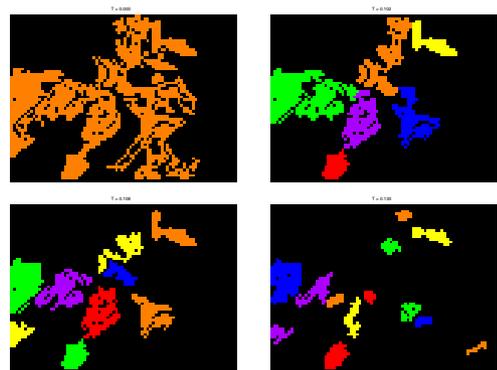


Figure 6: Results using the mouse sequence at different  $T$ .

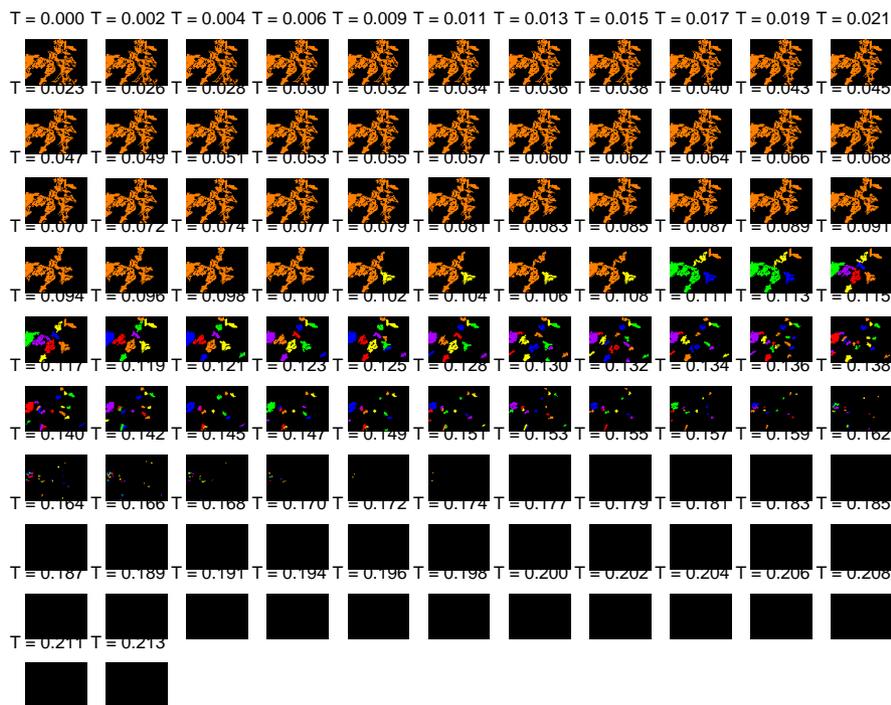


Figure 8: Mouse sequence results computing the parameter  $a$  using the *median*

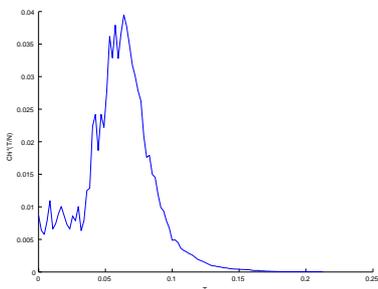


Figure 7: Mouse data: The susceptibility density of the clustering given different temperatures  $\chi T/N$  vs.  $T$ , the parameter  $a$  is computed using the *mean*.

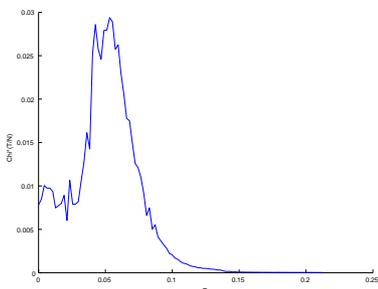


Figure 9: The susceptibility density of the clustering given different temperatures  $\chi T/N$  vs.  $T$ , the parameter  $a$  computed using the *median*.



Figure 10: Synthetic color image data.

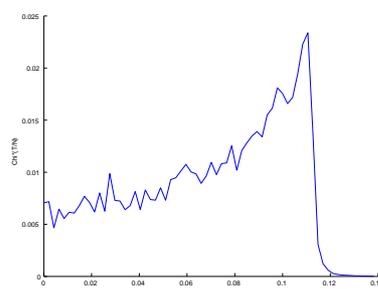


Figure 11: Color image data: The susceptibility density of the clustering given different temperatures  $\chi T/N$  vs.  $T$ , the parameter  $a$  computed using the *median*.

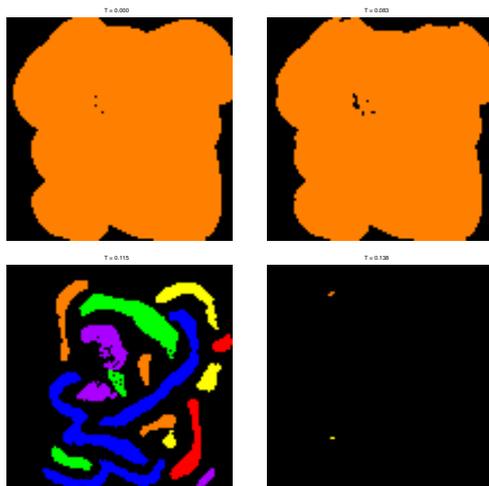


Figure 12: Results using the synthetic color data at different T.

- [4] C. Cédras and M. Shah. Motion-based recognition: A survey. <http://www.cs.ucf.edu/vision/publications.html>, 1994.
- [5] C. Fermler, D. Shulman, and Y. Aloimonos. The statistical of optical flow. *Computer Vision and Image Understanding*, 82:1–32, 2001.
- [6] H. Sawhney, Guo Y., and R. Kumar. Independent motion detection in 3d scenes. *IEEE Trans. on Pattern Recognition and Machine Intelligence*, 22(10):1191–1199, 2000.
- [7] D. Murray and B. Buxton. Scene segmentation from visual motion using global optimization. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 9, 1987.
- [8] J. Wang and E. Aldeson. Representating moving images with layers. *IEEE Trans. Image Processing*, 3, 1994.
- [9] N. Vasconcelos and A. Lippman. Empirical bayesian motion segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(2):217–221, 2001.
- [10] G. Celeux, F. Forbes, and N. Peyrard. EM procedures using mean field-like approximations for markov model-based image segmentation. *Pattern Recognition*, 36:131–144, 2003.
- [11] D. Cremers and C. Schnörr. Motion competition: Variational integration of motion segmentation and shape regularization. *Pattern Recognition(L. van Gool, editor) of LNCS*, 2449:472–480, 2002.
- [12] P. H. S. Torr and A. Zisserman. Concerning Bayesian motion segmentation, model averaging, matching and the trifocal tensor. *5th European Conference on Computer Vision*, 1406:511–527, 1998.
- [13] Toru Tamaki, Tsuyoshi Yamamura, and Noboru Ohnishi. Extracting human limb region using optical flow and nonlinear optimization. In *Proc. of ACCV2002*, pages 593–597, 2002.
- [14] A. Mansouri, B. Sirivong, and J. Konrad. Multiple motion segmentation with level sets. In *SPIE Image and Video Communications and Process*, volume 3974, pages 584–595, 2000.
- [15] J. Suri and S. Laxminarayan. *PDE and Level Sets: Algorithmic Approaches to Static and Motion Imagery (Topics in Biomedical Engineering)*. Kluwer Academic, 2002.
- [16] F. Campbell and Robson J. Application of fourier analysis to the visibility of gratings. *Journal of Physiology*, 197:551–556, 1968.
- [17] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.
- [18] D.J. Fleet and A.D. Jepson. Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5:77–104, 1990.
- [19] J.L Barron, A.D. Jepson, and J.K. Tsotsos. The feasibility of motion and structure from noisy time-varying image velocity information. *International Journal of Computer Vision*, 5:239–269, 1990.
- [20] S. Wiseman, M. Blatt, and Domany E. Superparamagnetic clustering of data. *Physics Review E*, 57:3767–3783, 1998.
- [21] R. Kindermann and J. Snell. *Markov Random Fields and Their Applications*. American Mathematical Society, 1980.
- [22] Wu Y. N., S. C. Zhu, and X. Liu. Equivalence of Julesz ensembles and frame models. *International Journal of Computer vision*, 43(3):245–264, 2001.
- [23] R. Swendsen and J-S. Wang. Nonuniversal critical dynamics in monte carlo simulation. *Physics Review Letters*, 58(2):86–88, 1987.
- [24] J-S Wang. Cluster monte carlo algorithms and their applications. *Recent Developments in Computer Vision, S.Z. Li et.al. (eds.)*, LNCS 1035:307–, 1996.
- [25] J. Punzicha and J. Buhmann. Multiscale annealing for grouping and unsupervised texture segmentation. *Computer Vision and Image Understanding*, 76(3):213–230, 1999.
- [26] G. Gilmer’farb. *Image Texture and Gibbs Random Fields*. Kluwer Academic Publishers, 1999.
- [27] Siu-Leong Iu. Robust estimation of motion vector fields with discontinuity and occlusion using local outliers rejection. *Journal of Visual Communication and Image Representation*, 6(2):132–141, 1995.