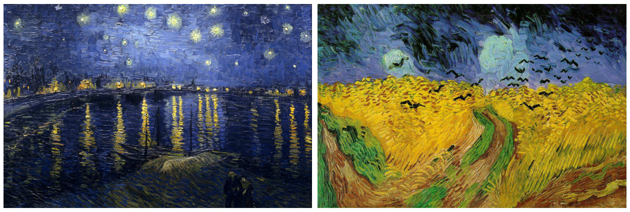

Figure 2: Van Gogh's mood. Left: "Starry Night over the Rhone". The top three predicted categories by our model are casual, modern, and chic. Right: "Wheat field". The top three predicted categories are casual, dapper, chic.

and then remove each kind of features out of the PFG model and evaluate the performance reduction (shown in Table 3). We can see that all the features have contributions to infer affects. Furthermore, HSV feature is most useful in terms of *Precision* which achieves a 8.88% reduction, while saturation and brightness features are the most helpful features in terms of *Recall* and *F1-Measure*. More importantly, the analysis confirms that the network information is one of the most important features. Without considering the network information, the F1-Measure drops 6.55%.

Effect of the size of training set in initialization. Inference accuracy depends on the size of training set for the initialization. A small number might result in high precision but low recall, while a large number might mean higher recall but would hurt the precision. Figure 1 shows how the average performance changes by varying the size of the training set. When the size of the training set is larger than 16,000, *Precision* grows slowly, which indicates the rationality of using 20,000 images as training set in our experiments.

4.3 Demonstrations

Inferring van Gogh's mood. We choose two typical van Gogh's paintings: "Starry Night over the Rhone" (Figure 2 left) and "Wheatfield with Crows" (Figure 2 right). We use our proposed model to infer van Gogh's mood from these two paintings. The results reflect the common affective cognition on these two paintings. For the first painting, the top three prediction categories are casual (probability: 19.3%), modern (15.03%), and chic (10.94%). The comments from *vangoghgallery* on this painting is quiet, rational, and stylish, which are all included in the semantic concepts of these three categories. For the second painting, the top three prediction categories are casual (19.09%), dapper (13.17%), chic (12.33%). The semantic concepts in casual and chic seem to contradict with the semantic concepts in dapper, such as carefree vs. awe-inspiring, or animated vs. bleak. This is the last painting before van Gogh's death, the comments from *vangoghgallery* on this painting is heavy, gloom, and insecure. These results indicate that our prediction results are almost consistent with the comments.

Inferring affects around special events. Here we download images from Flickr around Thanksgiving 2011, and use our model to predict the affective category of each image. Figure 3 shows affective distributions before and during Thanksgiving, with each containing 7,354 and 7,132 images respectively. We visualize the prediction results in Kobayashi's Color Image Scale. For each category, we use the typical five dominant colors to colorize its corresponding circle. And the area of a circle indicates the total number of images belong to this category. Before Thanksgiving, the public affects distribute normally in each category. During Thanksgiving Holiday, the public affects significantly concentrate on casual, the

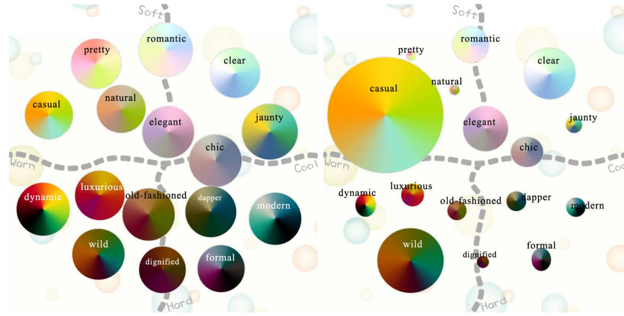

Figure 3: Affects around Thanksgiving 2011 inferred by images on Flickr. Left: affective distribution before Thanksgiving; Right: affective distribution after Thanksgiving. During Thanksgiving Holiday, the detected public affects are mainly on casual, which indicate cheerful, happy, free, and friendly.

semantic concepts of which are cheerful, happy, free, friendly, enjoyable, etc. That is a quite interesting but rational result.

5. CONCLUSION

In this paper, we study the problem of inferring affects from images in social networks. We first experimentally analyze features of color compositions that reflect the human's affects. And then, we propose a partially-labeled factor graph (PFG) model for inferring affective information on large scale images in social networks. Experiments demonstrate the effectiveness of the proposed model. As to the future work, we are planning to incorporate other features such as shapes for further improving the inferring accuracy.

Acknowledgements. This work is supported by the National Basic Research Program (973 Program) of China(2011CB302201), and Science Foundation of China (61003094, 90820304, 61073073), and also funded by Tsinghua National Laboratory for Information Science and Technology(TNList)Cross-discipline Foundation. Jie Tang is supported by Chinese National Key Foundation Research (No. 60933013) and National Significant Science and Technology Projects of China under Grant(No.2011ZX01042-001-002).

6. REFERENCES

[1] B. Frey and D. Dueck. Mixture modeling by affinity propagation. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *NIPS*, pages 379–386, 2006.

[2] S. Kobayashi. *Art of Color Combinations*. Kodansha International, 1995.

[3] J. Machajdik and A. Hanbury. Affective image classification using features inspired by psychology and art theory. In *Multimedia*, pages 83–92. ACM, 2010.

[4] L. Ou, M. Luo, A. Woodcock, and A. Wright. A study of colour emotion and colour preference. *Color Research & Application*, 29(3&4):232–240&292–298, 2004.

[5] Y. Shin and E. Kim. Affective prediction in photographic images using probabilistic affective model. In *ICIVR*, pages 390–397. ACM, 2010.

[6] M. Solli and R. Lenz. Color semantics for image indexing. In *ECCGIV*, 2010.

[7] J. Tanaka, D. Weiskopf, and P. Williams. The role of color in high-level vision. *Trends in cognitive sciences*, 5(5):211–215, 2001.

[8] W. Tang, H. Zhuang, and J. Tang. Learning to infer social ties in large networks. In *ECML/PKDD'11*, pages 381–397, 2011.