COLORIZING SINGLE-BAND THERMAL NIGHT VISION IMAGES

Xiaojing Gu[1, 2], Henry Leung[2], Shaoyuan Sun[1, 2] Jianán Fang[1] and Haitao Zhao[3]

College of Information Science and Technology, Donghua University[1],
Shanghai, China 201620
Department of Electrical and Computer Engineering, University of Calgary[2],
Calgary, Alberta, Canada T2N IN4
School of Aeronautics and Astronautics, Shanghai Jiao Tong University[3],
Shanghai, China 200240

ABSTRACT

We consider the problem of assigning single-band thermal night vision image with natural day-time color appearance automatically. We present an approach in which supervised learning is first used to estimate colors of monochromic images. Modeling color distribution of thermal imagery is a challenging problem, since there are insufficient local features for estimating the chromatic value at a point. Our model uses a statistical learning algorithm that incorporates multi-scale and spatially arranged image features, and it can be trained on a data set that contains thermal image and registered day-time color image pairs. Experimental results show that our approach leads to relatively accurate description of the desired color distribution and results in thermal images that appear smooth and natural color details, so that the overall scene recognition and situational awareness can be improved.

Index Terms— thermal imagery, color enhancement, supervised learning, night vision

1. INTRODUCTION

Thermal or longwave infrared imagery, which obtained from infrared sensor and can display invisible thermal energy emitted from objects, is a crucial source of information for sensor surveillance, reconnaissance, and security applications. Until recently the standard representation of a thermal night vision image is monochrome such as grayscale still. However, with the deeper reorganization on the role of color in human visual perception, there is a growing interest in displaying night vision imagery with colors.

In general, color image has obvious superiority over monochrome image on the visual tasks involving scene understanding and object detection by allowing the use of color perception and color knowledge. Experimental evidence has indicated that rendering night vision imagery with a natural color appearance that resembles its day-time looking may significantly improve the observer performance, achieving better recognition accuracy and reducing reaction time [1]; while an inappropriate false color appearance, which inconsistent with the surface colors in day light, could also be detrimental to human performance, due to the negation of observer's color knowledge from prior experience [2]. Thus, the goal of colorizing night vision images should be achieving a natural and day-time like color appearance so that the overall scene recognition and situational awareness can be improved.

Most works on displaying night vision imagery with natural colors have focused on "multi-band night vision image fusion" [3-7]. They generate colors depending on the complementary information from heterogeneous sensors. In this paper we consider the problem of automatically giving single-band thermal night vision imagery a natural day-time color appearance. We present a novel approach that combines machine learning and computer vision. Our motivation is twofold: First, since multi-band night vision system is much more complicated, expensive and less easy to carry than single-band, infrared sensor is often used separately and gives the most important target detection information (such as in the cases of surveillance systems and assistant driving systems); thus we consider it an important open problem to develop methods for colorizing thermal night vision imagery by analyses of itself, rather than by multi-band fusion. Second, although thermal imagery has very few textures, with low contrast, blurry details and narrow gray range, we believe that it is still possible for pixel distinction by incorporating the overall organization of the image form a "contextual" view, thereby chromatic value of each pixel can be estimated properly.

To our knowledge, there is little work on colorizing single-band thermal night vision imagery with natural colors. We propose an algorithm that learns color distribution model of thermal imagery using multi-scale and spatially arranged image features. Our model is based on statistical supervised learning methods such as linear regression, and can be trained on the combination of infrared image and corresponding natural color reference image. After training, our model could show a relatively accurate description of the desired color distribution.

2. RELATED WORK

Recently, there are several algorithms have been proposed for displaying night vision image with natural colors. Most of them focused on multi-band image fusion. [3] proposed an opponent processing color fusion method. Multi-band image was first fused into a false color image in RGB color space and then a color remapping was carried out in HSV color space in order to get a more natural color presentation. [4] and [6] demonstrated a new method based on "color transfer technique", a linear mapping of statistical properties was applied on the fused false color image in order to match those of a natural color reference image. More recently, [7] and [8] proposed a more practical method using a color look-up table that stored the day-time color values of a range of materials that might appear in the night vision scenes, and the look-up table was indexed by the combination of multi-band outputs.

Besides, Toet [5] had presented a method to colorize single-band intensified night vision image (low-light-level image). However, it required that the local features of night vision image such as mean and standard deviation of the luminance component must resemble those of the color reference image, thus this method is difficult to apply to thermal image, which with a very different luminance distribution from daylight image and lack of details.

3. METHOD

We apply supervised learning to predict the chromatic values as a function of the image. The general problem of adding color to grayscale image is a mapping from one-dimension to three-dimension which with uncertain solution, thus each grayscale value may correspond to a wide range of colors. In our approach we constrain this problem by "contextual information" of image, rather than "local information", since there are insufficient local features on thermal imagery such as textures and discriminative luminance values for predicting chromatic value at a point. We apply a multi-scale and spatially arranged vector capturing contextual features of each pixel, and then we use linear regression on these features to estimate the colors.

3.1. Feature Vector

Each pixel is measured by a set of multi-scale and spatially arranged neighborhoods around it. To capture contextual features, a small neighborhood covering the pixel as well as its four neighboring neighborhoods is considered and this is repeated at each of the three scales (the larger scale is containing the smaller scale as its central part and also covering surrounding areas around the smaller scale). Therefore, the feature vector of each pixel involves the information of itself, its direct neighborhood and its neighboring neighborhoods, which are from all the three scales, as show in Fig.1.

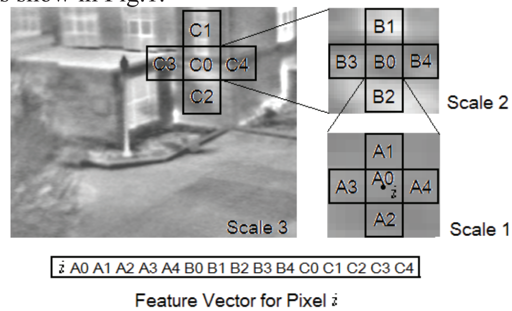


Figure 1. the feature vector for pixel i, which involves itself and its neighbors' information from all the three scales.

For each neighborhood of pixel i at scale 2 and scale 3 (the larger scales), we apply Law's masks [9] to measure texture energies, while for each neighborhood at scale 1 (the smallest scale) we only compute the mean and standard deviation of the luminance values over it, since there is little texture information at the smallest scale.

Law's masks is a powerful approach to texture analysis, it is simple, fast and high generalization. The basic nine masks are derived from three one-dimensional vectors L_3, E_3 and S_3, respectively, describe the following features: level, edge and spot. The one-dimensional vectors are as follows:

$$L_3 = (1 \quad 2 \quad 1); \quad \rightarrow \quad \text{Level detection}$$
$$E_3 = (-1 \quad 0 \quad 1); \quad \rightarrow \quad \text{Edge detection}$$
$$S_3 = (-1 \quad 2 \quad -1) \quad \rightarrow \quad \text{Spot detection}$$

Three 1-D Laws' vectors

By multiplying any vertical one-dimensional vector with a horizontal one, the nine two-dimensional masks $M_1,...,M_9$ of size 3×3 are constructed as:

$$\begin{matrix} L_3^T L_3 & L_3^T E_3 & L_3^T S_3 \\ E_3^T L_3 & E_3^T E_3 & E_3^T S_3 \\ S_3^T L_3 & S_3^T E_3 & S_3^T S_3 \end{matrix}$$

Nine 2-D Law's masks M_n

We convolute the thermal image $I(x,y)$ with each of these masks to obtain

$$T_n(x,y) = I(x,y) * M_n(x,y), \qquad n = 1,....,9 \tag{1}$$

Then, for pixel i, we estimate the texture energies En of its each of the five neighborhoods $N^i(p)$ at both scale 2 and scale 3 as:

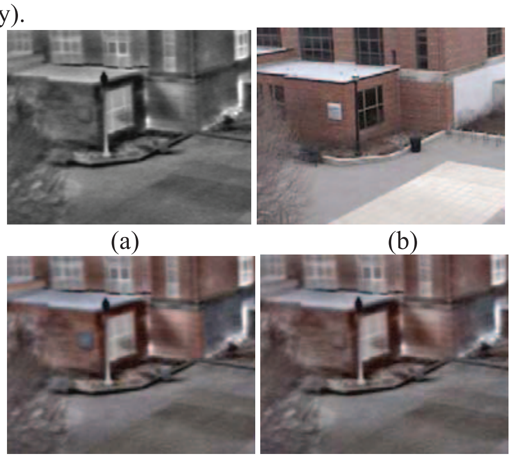
$$En_{N^i(p)}^{Scale}(i) = \sum_{x,y\in N^i(p)} \left|T_n(x,y)\right|, \qquad n = 1,...,9, \;\; p = 0,...,4, \;\; Scale = 2,3. \tag{2}$$

This gives $9\times5\times2 = 90$ feature coefficients for each pixel. Meanwhile, we also compute the mean and standard deviation of the luminance values over each neighborhood of pixel i at scale 1 and this gives another $2\times5 = 10$ feature coefficients. Including the luminance value of pixel

i itself, we totally get 101 dimensional feature vector for estimating color of a particular pixel.

3.2. Training

We create the labeled data from the combination of thermal image (Fig.2 (a)) and corresponding day-time color image with the registered scene (Fig.2 (b)), which means both the infrared sensor values and the corresponding natural colors are known. We present chromatic values in $L^*u^*v^*$ color space, because in this color space the perceived color differences correspond to Euclidean distances, which obviously facilitate color estimation problem in our approach. A fused image for training is formed by transforming color image form RGB into the $L^*u^*v^*$ color space and then replacing the original luminance channel L^* with its corresponding grayscale thermal image as shown in (Fig.2 (c)) (the result is transferred back to RGB for display).



(a) (b) (c) (d)

Figure 2. (a) thermal image. (b) day-time color reference image. (c) fused image by replacing L^* channel of (b) with (a). (d) coloring result of (a) by using the linear regression model that learned from (c).

Using the fused image and the feature vector for each pixel on the grayscale thermal image (L^* channel in fused image) as described above, we train a statistical linear model to estimate chromatic value (u^* and v^*-value) for each pixel. In the simplest case, a linear regression is used. Given training samples $\{\mathbf{f}_i, \mathbf{c}_i\}$ for $i \in G(x,y)$, where $G(x,y)$ is the fused image, $\mathbf{f}_i$ is the feature vector of pixel i on L^* channel of $G(x,y)$, $\mathbf{c}_i$ is the two-component color vector of pixel i that consisting of u^* and v^*-value of $L^*u^*v^*$ color space. The linear regression output value $\hat{\mathbf{c}}$ corresponding to a feature vector $\mathbf{f}$ is $\hat{\mathbf{c}} = \mathbf{a}^T\mathbf{f} + \mathbf{b}$, where $\mathbf{a}$ and $\mathbf{b}$ minimize the least-squares error:

$$arg\min \sum_{i \in G(x,y)} (\left\| \mathbf{a}^T\mathbf{f}_i + \mathbf{b} - \mathbf{c}_i \right\|_2)^2 \quad (3)$$

Fig.2 (d) shows the colorizing result by using this linear regression model on thermal image Fig.2 (a) after trained with Fig.2 (c). It can be seen that linear regression achieves a relatively accurate recovery of the ground truth color distribution. Furthermore, this model can also be trained with several pairs of thermal image and color reference image together, which leads to a more general application.

4. RESULTS

To test the coloring method for thermal imagery we use two couples of thermal image and its registered day-time color reference image as training data. These scenes, which are shown in Figs.3-4 (a) and Figs.3-4 (b), display many different components (buildings, grass, trees, pavements, and people). Fused images can be simply obtained by transforming color reference imagery (Figs.3-4 (b)) form RGB into the $L^*u^*v^*$ color space and then replacing the original luminance channel L^* with its corresponding grayscale thermal imagery (Figs.3-4 (a)), which are not shown here. The testing images are shown in Figs.3-4 (c), which with scenes that relevant to the training images. Both the training and testing images are obtained from OSU color-thermal Database [10]. All the following color results are transferred to RGB for display.

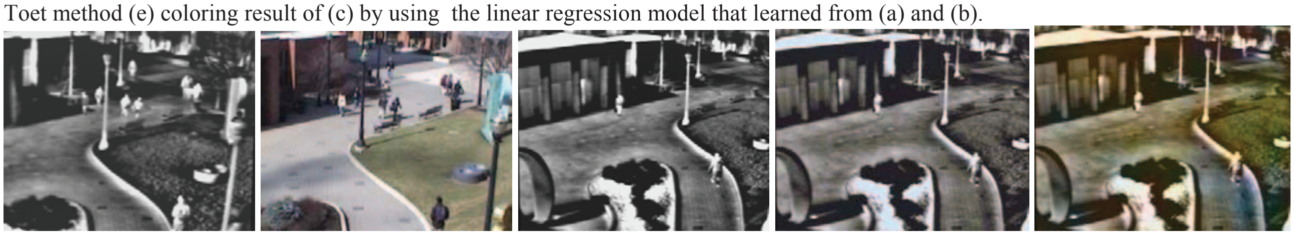
Figs.3-4 (d) show the coloring results of testing thermal images (Figs.3-4 (c)) by using the method presented in [5], applying Figs.3-4 (b) as color reference images respectively. Figs.3-4 (e) show the coloring results by using our linear regression model that learned from the fused image of Figs.3 (a)-(b) and Figs.4 (a)-(b), respectively. The results show that only employ the local features in [5] cannot achieve satisfied color because of the difficulty of pixel distinction on thermal images. While in our approach thermal images can be assigned natural color appearances that similar to the day-time color images, which enable fast scene recognition and interpretation by involving diagnostic colors (e.g. grass is green, building is red, different objects have different colors).

Figs.5 shows the colorized results of a linear regression model that trained with the fused images of Figs.3 (a)-(b) and Figs.4 (a)-(b) together. It can be seen that both different scenes (Figs.5 (a) and (c) depict similar scenes, and so does Figs.5 (b) and (d)) are assigned relatively proper color distribution. The results indicate that more general application can be achieved by training with several different scenes together.

In practice our approach will be used during the night when no corresponding daytime color images are available. The idea is to train the model in advance with a combination of thermal and color imagery of a scene or several scenes together, and then at night a suitable color scheme can be applied to the single-band thermal images with relevant scenes, which is very easy to be realized in surveillance systems and assistant driving systems.

(a) (b) (c) (d) (e)

Figure 3. (a) training thermal image. (b) training day-time color reference image. (c) testing thermal image. (d) coloring result by using Toet method (e) coloring result of (c) by using the linear regression model that learned from (a) and (b).



(a) (b) (c) (d) (e)

Figure 4. (a) training thermal image. (b) training day-time color reference image. (c) testing thermal image. (d) coloring result by using Toet method (e) coloring result of (c) by using the linear regression model that learned from (a) and (b).

(a) (b) (c) (d)

Figure 5. (a), (b), (c) and (d) are the coloring results of Fig.3 (a), Fig.4 (a), Fig.3 (c) and Fig.4 (c) respectively, using a linear regression model that trained with Figs.3 (a)-(b) and Figs.4 (a)-(b) together.

5. CONCLUSION

We have presented a framework for giving single-band thermal night vision image a natural day-time color appearance, which enable fast scene recognition and interpretation. Our model uses multi-scale and spatially arranged image features, and it can be trained on the combination of thermal image and corresponding color reference image. This approach can also apply to other colorizing tasks which dealing with grayscale image without fine textures. Our further experiments demonstrate that nonlinear model such as support vector regression can give a more accurate description of color distribution.

Acknowledgement

This work is supported by National Natural Science Foundation of China (No.60502042, 60705006, 60874113), Shanghai Rising-Star Program (No.06QA14003), the Research Fund for the Doctoral Program of Higher Education (No. 200802550007), the Key Creative Project of Shanghai Education Community (No.09ZZ66), and the Key Foundation Project of Shanghai (No.09JC1400700).

6. REFERENCES

[1] E.A. Essock, M.J. Sinai, J.S. McCarley, W.K. Krebs, J.K. DeFord, “Perceptual ability with real-world nighttime scenes: image intensified, infrared, and fused-color imagery,” Human Factors, 41(3), 438–452, 1999.

[2] M.J. Sinai, J.S. McCarley, W.K. Krebs, “Scene recognition with infra-red, low-light, and sensor fused imagery,” Proceedings of the IRIS Specialty Groups on Passive Sensors, 1-9, 1999.

[3] A.M. Waxman, A.N. Gove, D.A. Fay, J.P. Racamato, J.E. Carrick, M.C. Seibert, E.D. Savoye, “Color night vision: opponent processing in the fusion of visible and IR imagery,” Neural Networks, 10 (1), 1–6, 1997.

[4] A. Toet, "Natural colour mapping for multiband nightvision imagery," Information Fusion, 4, 155-166, 2003.

[5] A. Toet, "Colorizing single band intensified nightvision images," Displays, 26(1), 15-21, 2005.

[6] Y. Zheng and E. A. Essock, "A local-coloring method for night-vision colorization utilizing image analysis and fusion," Information Fusion, 9(2), 186-199, 2008.

[7] M.A. Hogervorst and A. Toet, "Fast and true-to-life application of daytime colors to night-time imagery," 10th International Conference on Information Fusion, 2007.

[8] A. Toet and M.A. Hogervorst, " Towards an Optimal Color Representation for Multiband Nightvision Systems," 12th International Conference on Information Fusion, 2009.

[9] Davies, E, *Machine vision: Theory, algorithms, practicalities* 2nd ed, Academic Press, 1997.

[10] J. Davis and V. Sharma, "Background-Subtraction using Contour-based Fusion of Thermal and Visible Imagery," Computer Vision and Image Understanding, 106(2-3), 162-182, 2007.