

基于小波图像融合的表情细节合成

王晓慧 贾 珈 蔡莲红

(普适计算教育部重点实验室(清华大学) 北京 100084)

(清华信息科学与技术国家实验室(筹) 北京 100084)

(清华大学计算机科学与技术系 北京 100084)

(wangxh09@mails. tsinghua. edu. cn)

Expression Detail Synthesis Based on Wavelet-Based Image Fusion

Wang Xiaohui, Jia Jia, and Cai Lianhong

(Key Laboratory of Pervasive Computing (Tsinghua University), Ministry of Education, Beijing 100084)

(Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing 100084)

(Department of Computer Science and Technology, Tsinghua University, Beijing 100084)

Abstract Expression details are texture changes caused by facial expression, such as wrinkles in the corner of the mouth when smiling and wrinkles on the forehead when surprising. Expression details can help to enhance the realistic experience of synthesized face image. In this paper, we propose to synthesize expression details by using the method of wavelet-based image fusion. We try to mine the texture feature of expression details for natural expression generation. In order to meet the requirements of individual expression detail synthesis, we use different wavelet transforms, such as the traditional wavelet transform and dual-tree complex wavelet transform, and kinds of fusion operators to get rich results. To seamlessly integrate the synthesized image of expression details to the output expressive face image, we select the optimal replacement for both images by clustering and graph cut method. Our proposed approach is applied to not only grayscale images, but also color images by the color space conversion. The experimental results show that the proposed method is effective on the expression detail synthesis, which can enhance the realistic experience of synthesized face image.

Key words wavelet transform; dual-tree complex wavelet transform (DT-CWT); image fusion; expression detail synthesis; optimal replacement selection

摘 要 表情细节是人脸表情变化时带来的皮肤纹理变化,表情细节合成有助于增强合成表情的真实感.提出了基于小波的图像融合方法,挖掘表情细节的纹理特征本质,并应用到表情细节的合成中,使合成的表情更加真实自然.为了满足表情细节个性化的需求,采用传统的小波变换和双树复小波变换,同时使用不同的融合算子,得到了丰富的表情合成.本方法不仅适用于灰度图像,通过颜色空间的转换还适用于彩色图像.最后还提出了基于聚类 and 图切割的最优替换区域选取方法,使得合成的表情细节区域与目标图像融为一体.

关键词 小波变换;双树复小波变换;图像融合;表情细节合成;最优替换区域选取

中图法分类号 TP391

表情细节是人在做表情时带来的皮肤纹理的变化.例如笑时嘴角会有皱纹;惊讶时会有抬头纹等.从二维图像^[1]到三维虚拟人^[2],表情合成一直是图像处理的研究热点.表情细节往往传递了丰富的“言外之意”,但是多数研究重点关注表情动作是否逼真,而忽略了表情细节,因此使得合成结果还是给人带来不自然的感觉.

表情细节合成的关键问题是:1)如何提取表情细节的本质,即纹理特征;2)如何将表情细节特征移植到目标图像中;3)由于表情细节因人而异,如何实现个性化表情的合成.

在表情细节合成研究中使用最广泛的是基于比率图(expression ratio image, ERI)^[3]的方法.ERI是由具有某一表情的图像和中性表情图片对应像素点的值相除得到的,可以是灰度的比率值或者是亮度的比率值等.表情细节的合成过程即在中性表情图像的基础上叠加 ERI.在表情比率图方法的基础上,姜大龙等人提出了局部表情比率图(PERI)的方法^[4].与基于 ERI 的合成方法不同的是 PERI 并不覆盖整个人脸,而是选择具有表情细节的局部区域,比如额头、嘴角等等进行合成计算.但是无论是 ERI 还是 PERI 方法都是单纯考虑了中性脸像和表情脸像的像素级差异,而不是表情细节的本质——纹理特征,所以效果都不尽理想.此外,以上方法依赖表情库的规模不能满足个性化表情细节合成的需求.

小波变换擅长提取纹理特征,在纹理移植方面的应用非常广泛^[5].因为小波特征能挖掘表情细节的纹理本质,同时图像融合方法可以实现表情细节图像模板到原始图像的无缝移植,所以本文提出了基于小波图像融合的表情细节合成方法,尝试将基于小波变换的图像融合方法应用到表情细节合成这一新的领域中,重点开展了以下 3 方面研究:1)通过使用传统的小波变换和双树复小波变换,采用不同的融合算子和改变小波分解层数的方法,得到了丰富的合成结果,从而不依赖于表情库就能满足表情细节个性化的需求;2)通过颜色空间的转换实现了彩色图像上的表情细节移植;3)表情细节的合成是对脸像局部区域进行的,本文提出了基于聚类和图切割的最优替换区域选取方法,使得合成的表情细节与目标图像融为一体.

1 基于小波的图像融合

1.1 小波变换与表情细节

小波(wavelet)具有良好的时频局部特性.从频域的角度看,人脸图像的低频分量反映出其固有的脸形与结构特征,而高频分量则承载着纹理上更细节的信息^[6],如皮肤褶皱就是人在做各种表情时所呈现出的细节特征.使用小波变换提取人脸纹理特征用于识别的很多,用于合成表情细节的很少.把人脸识别的方法用于人脸合成有利于发现其本质,使其合成的效果更符合自然真实的情况.

对图像进行二维离散小波变换,可以使用 Mallat 分解与重构算法^[7]来实现金字塔式的图像结构表示.例如,对图像 I 进行二维离散小波变换(2D DWT)的结果表示如下:

$$DWT(I, J) = (C^J(I), \mathbf{D}^J(I), \dots, \mathbf{D}^j(I), \dots, \mathbf{D}^1(I)), \quad (1)$$

其中, $\mathbf{D}^j(\cdot) = (D_h^j(\cdot), D_v^j(\cdot), D_d^j(\cdot))$, $1 \leq j \leq J$, 表示第 j 层 Mallat 分解得到的高频子图集,包括水平子图 D_h 、垂直子图 D_v 以及对角线方向子图 D_d ; $C^J(I)$ 表示 J 层分解后的低频子图.

1.2 图像融合策略

图像融合的定义是将两个或两个以上的图像利用一定的算法组合在一起形成一幅新的图像.本文基于小波的图像融合的基本思想是,首先对源图像进行小波多尺度分解,然后选择一定的融合算子得到融合后图像对应的小波系数,最后通过小波逆变换重构融合图像.

1.3 小波的选择

本文使用传统的小波变换和双树复小波变换进行小波多尺度分解.双树复小波变换(DT-CWT)是 1999 年由 Kingsbury^[8]提出的. DT-CWT 有两棵离散小波树,分别作用于图像的行和列,并行实现实部和虚部.每一级小波分解会得到 2 个低频子带和 6 个方向($\pm 15^\circ$, $\pm 45^\circ$, $\pm 75^\circ$)的高频子带.与传统的离散小波变换相比,DT-CWT 具有平移不变性、方向选择性、有限数据冗余等优点.通过 3.2 节的实验可以看出,当纹理方向比较多时,双树复小波可以得到更好的融合效果.

2 表情细节的合成

2.1 灰度图的表情细节合成

本文提出了将基于小波的图像融合技术应用到表情细节合成的方法,该方法原理如图 1 所示,其具体步骤如下:

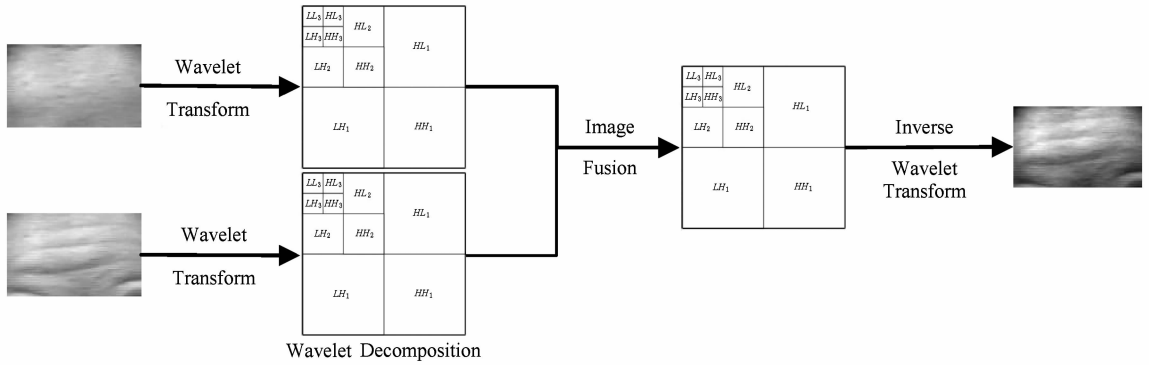


Fig. 1 Application framework.

图 1 应用原理图

对低频分量采用均值的融合算子;高频分量采用均值、绝对值最大和区域特性 3 种不同的融合算子. 均值的融合算子是待融合的两个分量的均值,即

$$D^j(I_F) = (D^j(I_s) + D^j(I_t))/2. \quad (2)$$

绝对值最大的融合算子是待融合的两个分量中绝对值最大的分量,即

$$D^j(I_F) = \max(D^j(I_s), D^j(I_t)). \quad (3)$$

基于区域特性融合规则^[9]是指根据待融合的两个分量所对应的图像局部区域的能量 E^j 和匹配程度 M^j , 设定一定的匹配度阈值 α , 最终得到融合算子:

$$E^j(m, n) = \sum_{m' \in K, n' \in L} \omega(m', n') [D^j(m + m', n + n')]^2; \quad (4)$$

$$M^j(m, n) = 2 \sum_{m' \in K, n' \in L} \omega(m', n') D_s^j(m + m', n + n') D_t^j(m + m', n + n') / (E_s^j + E_t^j); \quad (5)$$

$$\begin{cases} D^j(I_F) = D^j(I_s), & \text{当 } M^j < \alpha \ \& \ E_s^j \geq E_t^j; \\ D^j(I_F) = D^j(I_t), & \text{当 } M^j \geq \alpha \ \& \ E_s^j \geq E_t^j; \\ D^j(I_F) = W_{\max}^j D^j(I_s) + W_{\min}^j D^j(I_t), & \\ \quad \text{当 } M^j \geq \alpha \ \& \ E_s^j < E_t^j; \\ D^j(I_F) = W_{\min}^j D^j(I_s) + W_{\max}^j D^j(I_t), & \\ \quad \text{当 } M^j < \alpha \ \& \ E_s^j < E_t^j; \end{cases} \quad (6)$$

其中, $\begin{cases} W_{\min}^j = \frac{1}{2} - \frac{1}{2} \left(\frac{1 - M^j}{1 - \alpha} \right) \\ W_{\max}^j = 1 - W_{\min}^j \end{cases}$, K, L 定义了局部区域的大小, ω 是对应的权系数.

1) 对目标表情图像(待添加表情细节的人脸局部区域图像) I_t 和表情细节模板(含有表情细节的人脸局部区域图像) I_s 采用不同的小波函数进行小波多尺度变换, 分别得到低频和不同层级的高频分量 $D^j(I_s)$ 和 $D^j(I_t)$;

2) 对应分量采用不同的融合算子得到融合后相应的小波系数 $D^j(I_F)$.

3) 对最终低频和高频分量进行小波重构, 得到加入了表情细节的图像.

小波多尺度分解提取了表情细节的本质纹理特征; 图像融合技术实现了纹理的移植. 通过使用传统的小波变换和双树复小波变换, 采用不同的融合算子和改变小波分解层数的方法, 只使用一个表情细节模板就可以得到多种合成结果. 从而不依赖于表情库就能满足表情细节个性化的需求.

2.2 彩色图的表情细节合成

对于彩色图像, 表情细节区域既包括纹理信息也包括颜色信息. 为了使得合成的结果中保留其应有的肤色, 本文使用 HSI 颜色空间. 因为 HSI 空间中的 H 分量(Hue 色调)承载了图像的颜色类别信息, 可以从颜色空间中分离出来.

具体方法是先将目标表情图像和表情细节模板从 RGB 颜色空间转到 HSI 空间^[10], 然后分别对 S 分量(saturation 饱和度)和 I 分量(intensity 亮度)对应的图像进行融合处理, 将二者的融合结果和目标表情图像的 H 分量组合起来, 最后将其转换到 RGB 空间.

选取额头区域的两幅局部图像分别在 RGB 和 HSI 颜色空间进行融合, 其结果如图 2 所示.

从 3.2 节实验结果可以看出, RGB 空间的融合结果中包含了模板图像的肤色信息, 这样如果将其替换原人脸图像的额头部分, 存在肤色明显不同的

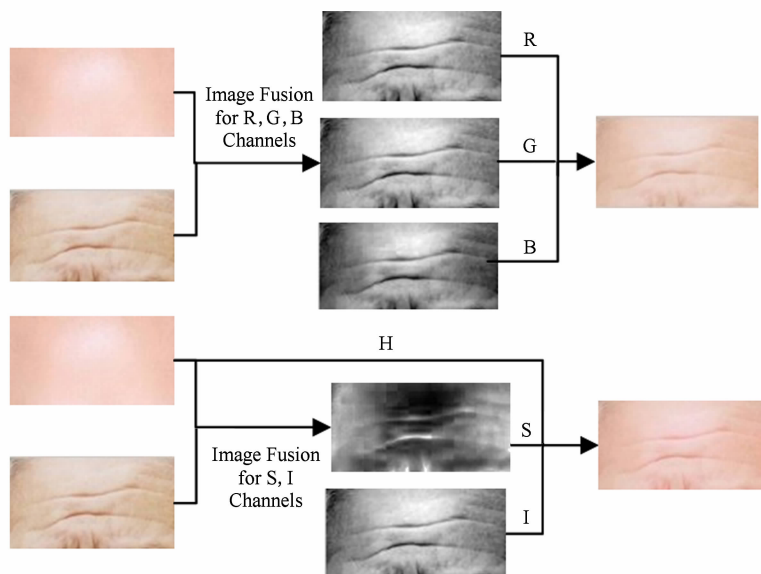


Fig. 2 Image fusion result in RGB and HIS color space.

图2 RGB和HSI空间的彩色图像融合结果对比

问题,使用 HSI 空间则克服了这个缺点.

2.3 基于聚类 and 图切割的最优替换区域选取

我们利用上述图像融合得到的表情细节区域图像替换目标图像中的相应区域,得到最终带有表情细节的人脸图像.为了使二者融为一体,看不出替换的边界,本文进一步提出了基于聚类和图切割相结合的最优替换区域选取方法.聚类采用 K -Means 算法.图切割(graph cut)^[11]的基本思想是根据合成的表情细节区域与目标图像所对应像素的差异构建一个有向图,然后使用最大流最小割理论,选择一个最优的不规则替换区域,使得最终合成的图像看不出替换边界.首先使用聚类算法把合成的表情细节区域的像素点进行聚类,个数较少的类中的像素点假设为边界上的像素;然后根据图切割算法去掉不是边界上的像素点,最终得到不规则的替换区域.以灰度图为例,使用图切割和聚类方法选取替换的边界,得到的结果如图 3 所示:



Fig. 3 Replacement selection results comparison.

图3 替换区域选取效果对比

图 3(a)所示为目标表情细节图像直接替代中性表情图像对应区域(矩形替换区域),可以明显看出替代边界;图 3(b)所示为用该方法从目标表情细节图像中选择一个不规则边界包围的区域,其结果自然很多.

3 合成实验及结果分析

本文中表情细节合成的方法用于表情合成的后处理.表情合成使用的是张申等人“基于语义维度的人脸表情生成”中的人脸模型和表情合成算法^[12],其输出带有表情但是缺少表情细节的人脸图像.该输出和从下面介绍的数据库中提取出来的表情细节模板作为本文表情细节合成方法的输入.实验首先介绍可以提取表情细节模板的数据库,然后分析基于小波图像融合方法的主要影响因素(包括小波函数、小波分解层数和融合算子)对图像融合性能的影响,最后实验示例直观展示表情细节的合成效果.

3.1 数据库

本文采用 Cohn-Kanade 人脸表情图片库作为实验数据^[13].该数据库中采集了 97 个大学生在表达 23 种不同面部动作时的表情图片序列.图 4 是数据库中的两幅图像,图 4(a)是中性表情图像,其额头区域、鼻尖到嘴角之间的区域等可以作为目标表情图像,图 4(b)是惊讶表情图像,其相应区域可以作为表情细节模板.通过 LAIV 实验室提供了 Cohn-Kanade 人脸图像库的 59 个特征点的人工标注结

果^[14],可以自动提取这些局部区域.

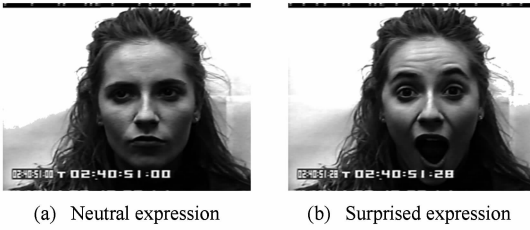


Fig. 4 Experimental image examples.

图 4 实验图像举例

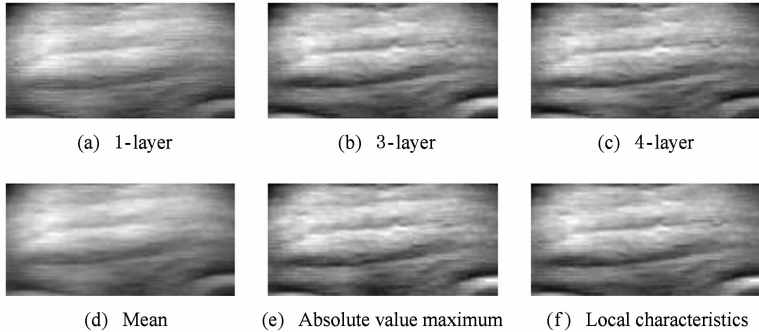


Fig. 5 Image fusion examples.

图 5 图像融合结果举例

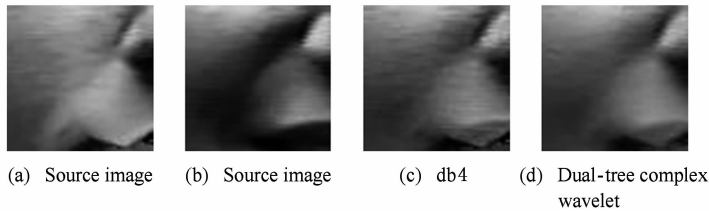


Fig. 6 Image fusion results using different wavelet function.

图 6 不同小波函数的融合结果比较

可以看出使用一个表情模板就能得到多种融合结果. 可以根据不同的应用场合来选择, 例如惊讶时年轻人的额头皱纹比较轻, 高频可以选择均值的融合算子; 老年人的皱纹比较重, 高频选择绝对值最大的融合算子. 这样不依赖表情库的规模就能满足个性化的表情细节需求.

3.3 实验分析

与传统的图像融合不同, 本文没有融合后的标准参考图像. 所以交叉熵、均方根误差、峰值信噪比等需要标准参考图像的融合效果的评价参量不适用本文. 本文选用熵(entropy)和互信息量 MI(mutual information)作为评价参量. 熵的大小反映了融合图像包含的平均信息量, 定义为

$$E = - \sum_{i=1}^L p_i \ln p_i, \quad (7)$$

3.2 实验结果

对数据库中图像的额头区域使用 db4 小波进行不同层数的分解和融合, 均值的融合算子得到的结果如图 5(a)~(c)所示. 使用 db4 小波进行 6 级小波变换, 高频使用不同的融合算子, 结果如图 5(d)~(f)所示.

选取表情为笑时具有多方向性的区域, 即鼻尖到嘴角之间的区域, 分别使用 db4 小波和双树复小波进行实验, 结果如图 6 所示.

L 是总灰度级, p_i 是归一化灰度直方图.

互信息量反映了融合图像从原始图像中提取的信息, 定义为

$$MI = \sum_{i=1}^L \sum_{j=1}^L \sum_{k=1}^L p_{stF}(i, j, k) \ln \frac{p_{stF}(i, j, k)}{p_{st}(i, j) p_F(k)}, \quad (8)$$

p_{st} 是图像 I_s 和 I_t 的归一化联合灰度直方图, p_{stF} 是源图像和融合图像的归一化联合灰度直方图. 熵和互信息量都是值越大融合图像所包含的信息越多.

3.3.1 小波函数及分解层数对图像融合性能的影响

对数据库中图像的额头区域, 分别使用 db4 小波和双树复小波进行不同层次的小波分解和融合, 高频融合算子取均值, 结果如表 1 所示. 小波分解层数越多, 熵和互信息量越大, 说明图像空间频带越丰富, 图像包含的细节信息越多. 同时实验表明层数

达到一定程度时,再增加分解层数效果不明显,而且增加计算量,所以一般在6层以下比较好.双树复小波的熵比db4小波大,但互信息量小,这解释了为什

么使用双树复小波融合的结果纹理更光滑,从3.2节的实验结果可以看出双树复小波方向选择性的优势.

Table 1 Performance under Different Wavelet Functions and Decomposition Levels

表 1 小波函数及分解层数对性能的影响

Fusion Operator	db4						DT-CWT					
	2	3	4	5	6	7	2	3	4	5	6	7
Entropy	4.3094	4.3075	4.3079	4.3082	4.3084	4.3085	4.5469	4.4506	4.5365	4.7167	4.9018	5.0027
MI	4.1769	4.2089	4.2502	4.2685	4.2680	4.2609	3.5099	3.5374	3.5901	3.7833	4.0000	4.1549

3.3.2 融合算子对图像融合性能的影响

对图像的额头区域,使用db4小波进行6层小波变换和融合,高频选择不同的融合算子,局部特性融合算子中区域设为 3×3 ,阈值为0.5,结果如表2所示:

Table 2 Performance under Different Fusion Operators

表 2 融合算子对性能的影响

Fusion Operator	Mean	Absolute Value	Local
		Maximum	Characteristics
Entropy	4.3084	4.7124	4.6434
MI	4.2680	3.7008	3.6369

从3.2节的实验结果中也可以看出融合算子为均值时,融合后图像平滑,皱纹痕迹不明显;融合算子为绝对值最大时,皱纹痕迹非常明显;融合算子为局部区域特征时,效果介于二者之间,皱纹痕迹明显且比较平滑.

4 结 论

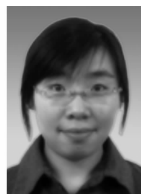
本文将基于小波的图像融合技术应用到表情细节合成中,通过改变小波分解的层数,采用传统的小波和双树复小波多种类型的小波函数,使用不同的融合算子,达到合成结果多样化的效果,满足了表情细节个性化的需求.通过颜色空间的转换,使得该方法不仅应用于灰度图还适用于彩色图像.最后提出了基于聚类 and 图切割相结合的最优替换区域选取方法,使得合成结果很好地融入到目标图像中.

参 考 文 献

- [1] Wu Juebo, Chi Hehua, Wang Shuliang, et al. Facial expression synthesis based on cloud model [C] //Proc of the 2nd Int Workshop Intelligent Systems and Applications (ISA). Piscataway, NJ: IEEE, 2010: 1-4
- [2] Bozkurt E, Erdem C, Erzin E, et al. Speech-driven automatic facial expression synthesis [C] //Proc of 3DTV Conf: The True Vision-Capture, Trans and Display of 3D Video. Piscataway, NJ: IEEE, 2008: 273-276
- [3] Liu Zicheng, Shan Ying, Zhang Zhengyou. Expressive expression mapping with ratio images [C] //Proc of Int Conf on Computer Graphics and Interactive Techniques. New York: ACM, 2001: 271-276
- [4] Jiang Dalong, Gao Wen, Wang Zhaoqi, et al. Realistic 3D facial animations with partial expression ratio image [J]. Chinese Journal of Computers, 2004, 27(6): 750-757 (in Chinese)
(姜大龙, 高文, 王兆其, 等. 面向纹理特征的真实感三维人脸动画方法[J]. 计算机学报, 2004, 27(6): 750-757)
- [5] Zhang Jianming, Fan Lijing, Zhan Yongzhao, et al. Reconstruction of human face for different age based on the Log-Gabor wavelet [C] //Proc of Int Symp on Information Science and Engineering. Piscataway, NJ: IEEE, 2008: 354-357
- [6] Zhao W, Chellappa R, Rosenfeld A, et al. Face recognition: A literature survey [J]. ACM Computing Surveys, 2003, 35(4): 399-458
- [7] Shan Ying, Liu Zicheng, Zhang Zhengyou. Image-based surface detail transfer [J]. IEEE Computer Graphics and Applications, 2004, 24(3): 30-35
- [8] Kingsbury N G. The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters [C] //Proc of the 8th IEEE Digital Signal Processing Workshop. Piscataway, NJ: IEEE, 1998: 86-89
- [9] Liu Guixi, Yang Wanhai. A wavelet-decomposition-based image fusion scheme and its performance evaluation [J]. Acta Automatica Sinica, 2002, 28(6): 927-934 (in Chinese)
(刘贵喜, 杨万海. 基于小波分解的图像融合方法及性能评价[J]. 自动化学报, 2002, 28(6): 927-934)
- [10] Rafael C G, Richard E W, Steven L E. Digital Image Processing Using MATLAB [M]. Translated by Ruan Qiuqi, et al. Beijing: Electronic Industry Press, 2005 (in Chinese)

(瓦萨雷斯,等. 数字图像处理[M]. 阮秋琦等,译. 北京: 电子工业出版社, 2005)

- [11] Kwatra V, Schodl A, Essa I, et al. Graphcut textures: Image and video synthesis using graph cuts [J]. ACM Trans on Graphics, 2003, 22(3): 277-286
- [12] Zhang Shen, Jia Jia, Wang Xiaohui, et al. Facial expression synthesis based on semantic dimensions [J]. Journal of Tsinghua University: Science and Technology, 2011, 51(1): 80-84 (in Chinese)
(张申, 贾珈, 王晓慧, 等. 基于语义维度的人脸表情生成 [J]. 清华大学学报: 自然科学版, 2011, 51(1): 80-84)
- [13] Kanade T, Cohn J, Tian Y. Comprehensive database for facial expression analysis [C] //Proc of the 4th IEEE Int Conf on Automatic Face and Gesture Recognition. Piscataway, NJ: IEEE, 2000: 46-53
- [14] Lipori G. Manual annotations of facial fiducial points on the Cohn Kanade database [OL]. [2009-11-10]. <http://lipori.dsi.unimi.it/download.html>.



Wang Xiaohui, born in 1987. PhD candidate of Tsinghua University. Her current research interests include image processing and pattern recognition.



Jia Jia, born in 1981. Assistant professor of Tsinghua University. Her main research interests include auditory perception computing and cognitive computing(jjia@tsinghua.edu.cn).



Cai Lianhong, born in 1945. Professor of Tsinghua University. Her main research interests are human computer speech interactive and affective computing(clh-dcs@tsinghua.edu.cn).

2013 年中国计算机学会人工智能会议(CCF AI2013)

征文通知

中国计算机学会人工智能会议由中国计算机学会主办,中国计算机学会人工智能与模式识别专业委员会协办,每两年召开一次. 本届会议将于 2013 年 7 月 31 日至 8 月 2 日在北京举行,会议由北京航空航天大学计算机学院承办. 本次会议旨在为中国人工智能领域的学者提供一个交流最新研究成果的平台,汇聚从事人工智能理论、技术及应用研究的人员,广泛开展学术交流,研讨发展战略,以期促进人工智能理论、技术及应用的发展.

会议录用的论文将推荐到《计算机研究与发展》、《模式识别与人工智能》、《计算机科学》、《计算机科学与探索》等国内权威或核心期刊上发表.

征文范围(但不限于)

- 人工智能理论基础
- 知识科学与知识工程
- 智能机器人与机器人足球
- 机器感知与虚拟现实
- 智能控制与智能管理
- 图像和语音处理
- 生物信息学与人工生命
- 机器学习
- 数据挖掘
- 社会网络分析及应用
- 神经网络与计算智能
- 人工免疫
- 粗糙集与软计算
- 时空知识表示、推理与挖掘
- 模式识别
- 人工智能应用
- 自然语言处理和机器翻译

征文要求

- 1) 论文应是未发表的研究成果.
- 2) 大会仅接收中文征文.
- 3) 论文应包括题目、作者信息、摘要、关键词、正文和参考文献,论文格式请参考《计算机研究与发展》的投稿格式.
- 4) 论文需附通讯作者的联系地址、电话或手机及 E-mail 地址.
- 5) 论文请通过大会网站注册的 EasyChair 系统提交,网址为 <https://www.easychair.org/conferences/?conf=ccfai2013>.

重要日期

征文截止日期: 2013 年 4 月 30 日

录用通知日期: 2013 年 5 月 30 日

联系方式

联系人: 孙 嘉

联系电话: 010-82316712

会务邮箱: ccfai2013@buaa.edu.cn

会议网址: <http://www.ccfai2013.org>