

基于支持向量聚类的多聚焦图像融合算法^{*}

张诚成^{1,2,3+}, 胡金春^{1,2,3}

¹(清华大学 计算机科学与技术系,北京 100084)

²(清华大学 智能技术与系统国家重点实验室,北京 100084)

³(清华信息科学与技术国家实验室,北京 100084)

Exploiting SVC Algorithm for Multifocus Image Fusion

ZHANG Cheng-Cheng^{1,2,3+}, HU Jin-Chun^{1,2,3}

¹(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

²(State Key Laboratory on Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China)

³(Tsinghua National Laboratory for Information Science and Technology, Beijing 100084, China)

+ Corresponding author: Phn: +86-10-62796711 ext 8042, E-mail: zhang-cc04@mails.tsinghua.edu.cn, http://www.csai.tsinghua.edu.cn/

Zhang CC, Hu JC. Exploiting SVC algorithm for multifocus image fusion. Journal of Software, 2007,18(10): 2445-2457. <http://www.jos.org.cn/1000-9825/18/2445.htm>

Abstract: This paper proposes a SVC (support vector clustering) based fusion rule according to unsupervised learning strategy. By employing the rule in multifocus image fusion applications, it solves the problems of region overlapping and abrupt transition brought about by the SVM (support vector machine) based fusion rule. The quality of the fused image is further enhanced. The undecimated discrete wavelet transform is applied to source images for multiresolution decomposition. Image feature data is extracted by means of grid, and it is then fed into the SVC algorithm which will generate distinct clusters. These resultant clusters are further processed by the domain discrimination algorithm and eventually distributed to two separate domains defined as complementary domain and redundant domain, in which choose-max method and weighted average method are used respectively to produce multiresolution representation of the fused image. Finally, the fused image is reconstructed by performing the corresponding inverse wavelet transform. The relation between the parameter q of SVC algorithm and the parameter $RMSE$ used to evaluate the fused image is studied in detail. It is indicated by theoretical analysis and experimental results that SVC is appropriate for image fusion. Moreover, comparative studies show that the proposed SVC based fusion rule outperforms the existing SVM based ones.

Key words: support vector clustering; support vector machine; image fusion; multiresolution analysis; fusion rule

摘要: 从无监督机器学习角度提出了一种基于 SVC(support vector clustering)的图像融合规则,解决了基于 SVM(support vector machine)的融合规则在处理多聚焦图像融合问题时所引起的区域混叠与非平滑过渡问题,进一步提高了融合图像的质量.使用非降采样离散小波变换对源图像进行多分辨率分解,基于网格提取源图像的特征图像特征集合作为 SVC 的输入数据集,聚类结果最终由区域鉴别算法分配到两个区域:互补信息区域和冗余信息区

* Supported by the National Natural Science Foundation of China under Grant No.10577012 (国家自然科学基金)

Received 2006-03-23; Accepted 2006-08-16

域,并分别采用选择法和加权平均法生成融合图像的多分辨率表示,通过对这一多分辨率表示进行小波逆变换重构融合图像.详细研究了 SVC 的参数 q 与融合效果的评价参数 $RMSE$ 之间的关系.理论分析及实验结果均表明,SVC 用于图像融合问题是合适的,而且比较实验显示,基于 SVC 的融合规则优于基于 SVM 的融合规则.

关键词: 支持向量聚类;支持向量机;图像融合;多分辨率分析;融合规则

中图法分类号: TP391 文献标识码: A

图像融合技术综合处理具有相同成像内容、不同成像形式的多幅源图像,目的是生成更符合人类视觉特性、更适合计算机处理的新的图像^[1],新图像所提供的信息是任何一幅源图像所无法独立提供的.图像融合技术已成为计算机视觉、遥感图像处理、医学图像处理、军事应用等领域的重要研究问题.

融合规则是图像融合技术的核心,也是至今仍有待探索的困难问题.自从Burt等人^[2]提出基于多分辨率框架的图像融合技术以来,多分辨率分析已经历了拉普拉斯金字塔、比率低通金字塔、梯度金字塔、形态学金字塔、离散小波变换、离散小波框架、可变方向多分辨率分析等技术阶段,其所能提取的图像特征越来越精细.与多分辨率分析所具备的越来越强的特征提取能力相比,融合规则对图像特征的利用效率仍然较低.Zhang等人^[3]总结了多分辨率图像融合技术中使用频率最高的两种融合规则——选择融合和加权平均融合,虽然它们具有计算简单的优点,但同时存在融合图像失真、对比度信息损失等问题.为了解决这些问题,人们从不同角度提出了更为复杂的融合规则,如多层次融合规则^[4]、基于神经网络的融合规则^[5,6]等,其目的均是为了更深入地挖掘图像特征中所蕴含的语义信息,提高特征利用效率.

Li等人^[7]从统计学习角度研究了多聚焦图像融合问题,通过训练支持向量机(support vector machine,简称SVM)实现对于源图像中清晰像素与模糊像素的分类,进而生成融合图像.这样的融合规则充分考虑了特征数据样本在数据空间中的分布情况及相互关系,较传统的融合规则更为合理.然而,基于SVM的融合规则存在两个问题:

第一,区域混叠问题.即融合图像不能完全再现源图像中的清晰目标,而在清晰目标中包含了相应模糊目标中部分像素所构成的区域.这样的区域是成块出现的,一致性校验(consistency verification)难以处理.直观上看,融合图像中出现模糊斑点,图像质量有所下降.通过观察 SVM 的输出可以发现,区域混叠问题产生的原因是 SVM 错误地判断了一部分像素的清晰程度,这与 SVM 的泛化能力有关.需要指出的是,清晰与模糊这两个概念本身难以精确定义,将同样一对目标图像置于不同的背景图像中,对于哪一幅目标图像更为清晰的判断可能截然相反.图 1 直观地显示了这两个概念的相对性.图 1 包含 4 幅图像,同一行的两幅图像具有完全相同的背景灰度分布,同一列的两幅图像具有完全相同的前景(字母)灰度分布.可以看出,在浅色背景下,深色字母显得更为清晰;而在深色背景下,浅色字母显得更为清晰.为抑制这种二义性对分类器泛化能力的影响,必须保证样本特征对于样本清晰程度的刻画足够准确.文献[7]设计的样本特征缺乏尺度自适应能力,难以有效解决二义性问题.而且从语义信息处理的角度讲,直接对像素分类不尽合理.与一般的分类问题相比,在清晰与模糊的分类器设计中,训练样本的构造更加困难.文献[7]并没有分析经过训练的 SVM 的泛化能力.



Fig.1 Illustration of the relativity of clearness and blurriness

图 1 清晰与模糊的相对性示例

第二,非平滑过渡问题.多聚焦图像融合并不是简单地将每幅源图像的清晰目标拼接在一起,必须对来自不同源图像的目标之间的过渡区域进行平滑处理.基于 SVM 的融合规则只将源图像的像素分为清晰与模糊两类,清晰像素被引入融合图像,模糊像素则被完全抛弃.这样的方案等价于清晰目标的简单拼接,在某些情况下,融

合图像会出现拼接痕迹。

针对基于 SVM 的融合规则在处理多聚焦图像融合时存在的两个问题,本文从无监督机器学习角度提出了基于支持向量聚类(support vector clustering,简称 SVC)的融合规则.首先使用 à trous 小波对源图像进行多分辨率分解,然后基于网格提取源图像的特征.图像特征组成的集合作为 SVC 的输入数据集,聚类结果最终由区域鉴别算法分为两个区域:互补信息区域和冗余信息区域.对于互补信息区域,采用选择法生成融合图像的多分辨率表示;对于冗余信息区域,采用加权平均法生成融合图像的多分辨率表示.最后对这一多分辨率表示进行小波逆变换,重构融合图像.由于聚类算法不需要训练样本集,所以基于聚类的融合规则对源图像的依赖性更小,对像素清晰程度的判别更准确,适用范围更广.同时,冗余信息区域的引入与处理考虑了不同目标之间的过渡问题.实验结果表明,基于 SVC 的融合规则有效地解决了区域混叠问题与非平滑过渡问题.

本文第 1 节简要介绍 SVC 算法.第 2 节详细给出基于 SVC 的图像融合技术,重点研究 SVC 在多聚焦图像融合中的应用特点以及融合规则的设计原理.第 3 节给出实验结果及分析.第 4 节总结全文.

1 SVC

Ben-Hur 等人^[8]在 SVM^[9]方法的基础上提出了 SVC 算法.对这一算法的研究工作在逐渐展开^[10-12].SVC 算法所能获得的聚类效果不受聚类形状限制,它也不需要预先给定聚类数目等参数.因此,它适用于对分布情况复杂的数据点聚类.SVC 算法首先使用高斯核函数 $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-q\|\mathbf{x}_i - \mathbf{x}_j\|^2)$ 将数据点从数据空间映射到高维特征空间,然后在特征空间中寻找能够包含所有数据点映像的最小超球.如果将最小超球面映射回数据空间,则它会形成一组轮廓线,这些轮廓线可以被解释为聚类的边界线,每一条轮廓线所包围的数据点属于同一聚类.基于以上原理,SVC 算法可以分为两部分:第 1 部分寻找最小超球,最终归结为求解一个二次规划问题;第 2 部分将各个数据点分配到相应的聚类中去.给定一对数据点,如果它们属于不同聚类,则连接两点的任意路径在特征空间的投影必穿过最小超球面.据此定义无向图 G 的邻接矩阵, G 的每一个连通分量即为一个聚类.可以使用 DFS(depth first search)算法求解图 G 的连通分量.SVC 的详细内容参阅文献[8].

2 基于 SVC 的图像融合算法

2.1 图像特征提取

本文研究两幅源图像的融合问题.设两幅源图像已经过配准,对于相同成像条件下的两个目标 O_1 和 O_2 ,源图像 A 中 O_1 清晰 O_2 模糊,源图像 B 中 O_1 模糊 O_2 清晰,融合算法的目的是获得图像 F ,使得 O_1 和 O_2 均清晰.SVC 用于划分源图像的互补信息区域和冗余信息区域,而所谓互补信息区域就是源图像中清晰程度变化较大的区域,冗余信息区域就是源图像中清晰程度变化较小的区域.为了正确划分这两个区域,SVC 需要能够度量源图像清晰程度的特征数据.本文选用 à trous 小波^[13,14]提取源图像特征,原因如下:首先,à trous 小波变换没有降采样过程,所以在变换过程中图像尺寸不发生变化,这使得各层系数对应关系明确,有利于生成特征向量;其次,à trous 小波具有平移不变性,这一特性可以在一定程度上抑制配准误差对融合结果的影响.

à trous 小波变换原理如下:对于图像 I_0 ,使用低通滤波器 Ft 对其进行 n 层滤波,每层滤波前 Ft 进行升采样.

$$I_1 = Ft(I_0), I_2 = Ft(I_1), \dots, I_n = Ft(I_{n-1}) \quad (1)$$

如此得到对比度逐渐降低的低频分量序列 I_1, I_2, \dots, I_n .高频分量 w_k 的计算方法为

$$w_k = I_{k-1} - I_k, k=1, 2, \dots, n \quad (2)$$

图像的重构可由相应的小波逆变换实现:

$$I_0 = \sum_{k=1}^n w_k + I_n \quad (3)$$

设源图像尺寸为 $r \times c$.考虑到多聚焦图像融合问题的特点,源图像的清晰部分与模糊部分一般是连续的区域,而不是离散的像素,所以,使用网格来分割源图像是合理的,同一网格内的所有像素被视为具有共同的性质.设网格尺寸为 $r_g \times c_g$,则特征提取方法如式(4)、式(5)所示.

$$f(i, j) = \begin{bmatrix} Gw_1^A(i, j) - Gw_1^B(i, j) \\ Gw_2^A(i, j) - Gw_2^B(i, j) \\ \vdots \\ Gw_n^A(i, j) - Gw_n^B(i, j) \end{bmatrix}, i=1, \dots, (r/r_g), j=1, \dots, (c/c_g) \quad (4)$$

$$Gw_k^I(i, j) = \sum_{a=1}^{r_g} \sum_{b=1}^{c_g} |w_k^I((i-1)r_g + a, (j-1)c_g + b)| \quad i=1, \dots, (r/r_g), j=1, \dots, (c/c_g), k=1, \dots, n, I=A, B \quad (5)$$

式(5)计算图像 I 的第 k 层清晰程度矩阵 Gw_k^I , 矩阵元素为图像 I 第 k 层高频分量 w_k^I 中对应网格内所有小波系数绝对值之和. 式(4)计算每个网格的 n 维特征向量, 该向量反映了两幅源图像中对应网格的清晰程度的差异. 由式(5)可知, Gw_k^I 的尺寸为 $(r/r_g) \times (c/c_g)$, 而 w_k^I 的尺寸为 $r \times c$. 可见, 基于网格的图像特征提取方法减少了特征数据量, 有利于进行支持向量聚类. f 由 $(r/r_g) \times (c/c_g)$ 个 n 维特征向量组成. 由式(4)可知, 如果与特征向量对应的 n 维空间中的数据点距离坐标系原点较远, 则相应的网格属于互补信息区域; 反之, 如果数据点距离坐标系原点较近, 则相应的网格属于冗余信息区域. 而对于远与近的判断需要通过聚类算法实现.

需要指出的是, 网格特征的准确程度依赖于目标的形状, 如果源图像中包含具有不规则几何形状的目标, 则互补信息区域与冗余信息区域交界处的网格特征是不准确的, 同一网格内一部分像素属于互补信息区域, 而其他像素属于冗余信息区域. 为抑制这种区域边界效应, 需要单独处理边界网格. 第 2.2 节首先提出互补信息区域与冗余信息区域的初始划分策略, 进而提出了多分辨率边界网格处理策略, 这两个策略构成完整的区域鉴别算法.

2.2 融合规则

经过特征提取, 多聚焦图像融合问题可以视为聚类问题. 理想情况下, 互补信息区域和冗余信息区域各成一类, 则多聚焦图像融合问题成为一个简单的问题. 然而, 实际的多聚焦图像融合问题并非如此.

令 SVC 的参数 $C=1$, 不考虑 BSV (bounded support vector), 则随着参数 q 的增大, SV (support vector) 的个数增多, 聚类数目增大. 对于互补信息区域, 从统计角度讲, 图像清晰程度变化量的均值较大, 方差一般也较大, 即: 区域内图像清晰程度的变化是不一致的, 互补信息区域可进一步划分为一系列子区域. 因此, 互补信息区域可以对应 n 维空间中的多个聚类. 另外, 从互补信息区域向冗余信息区域过渡, 图像的清晰程度一般是连续变化的, 这使得聚类成为一个两难问题. 根据 SVC 的特点, 如果 q 较小, 则聚类边界线分裂不充分, 互补信息区域的很多子区域被聚到冗余信息区域中去; 如果 q 较大, 则聚类精细, 互补信息区域的各个子区域被认为是独立的区域. 因此在 q 值给定的条件下, 互补信息区域和冗余信息区域的确定需要一套规则.

设 f 中欧氏范数最小的 n 维特征向量所属的聚类为 c_{\min} , 将 c_{\min} 的均值向量的欧氏范数作为阈值, 在所有聚类中, 均值向量的欧氏范数小于或等于该阈值的聚类被分配到冗余信息区域. 对于剩余的聚类, 首先根据均值向量各分量的变化方向将其中的一部分聚类分配到互补信息区域, 然后按照最小欧氏距离原则将另一部分聚类分配到上面已经确定的冗余信息区域和互补信息区域中去. 这样就实现了第 2.1 节中所要求的对于远与近的判断, 计算出了互补信息区域与冗余信息区域的初始划分.

初始划分一旦确定, 需要使用边界网格处理策略来调整这一划分, 以使互补信息区域的形状逼近图像中目标的真实形状. 本文提出多分辨率边界网格处理策略. 设 r_domain 表示冗余信息区域, c_domain_A, c_domain_B 表示互补信息区域, 其中, c_domain_A 表示源图像 A 清晰而源图像 B 模糊的对应区域, c_domain_B 表示源图像 B 清晰而源图像 A 模糊的对应区域. 首先使用 4-邻域搜索来定位边界网格, 即: 如果某一网格的上下左右 4 个相邻网格中至少存在一个与该网格所属区域不同的网格, 则将该网格标记为边界网格.

定义边界矩阵:

$$E = (e_{ij})_{(r/r_g) \times (c/c_g)}, e_{ij} = \begin{cases} 1, & \exists g(((g = f(i, j+1)) \vee (g = f(i, j-1)) \vee (g = f(i+1, j)) \\ & \vee (g = f(i-1, j))) \wedge (f(i, j) \in S_1) \wedge (g \in S_2) \wedge (S_1 \cap S_2 = \emptyset)). \\ 0, & \text{Otherwise} \end{cases}$$

其中, $S_1, S_2 \in \{r_domain, c_domain_A, c_domain_B\}$. 可知 $e_{ij}=1$ 表示 $f(i, j)$ 为边界网格. 边界网格特征是不准确的, 因而以所有非边界网格特征为依据计算互补信息区域与冗余信息区域的特征向量, 计算方法如下:

$$v_s = \frac{1}{N_s} \sum_{f(i, j) \in S, e_{ij}=0} f(i, j) \quad (6)$$

其中, $S \in \{r_domain, c_domain_A, c_domain_B\}$, N_s 表示区域 S 中非边界网格的数目.

根据特征相似性原则将边界网格重新分配到互补信息区域与冗余信息区域中去. 特征相似性原则的定义是: 如果某一边界网格距离一个区域充分近, 而距离其他区域充分远, 则允许将该边界网格分配到与之距离充分近的区域中去; 否则认为该网格与各个区域的相似性难以判断, 放弃分配. 基于多分辨率分析的思想, 如果无法准确判断某一边界网格所属的区域, 则将其行列尺寸减半, 一个边界网格被分为 4 个边界网格, 再根据特征相似性原则将这 4 个网格分配到互补信息区域与冗余信息区域中去. 这一过程迭代进行, 直到网格尺寸减为 1×1 , 即单像素网格. 对于仍无法判断所属区域的像素, 根据最小欧氏距离原则将其分配到相应区域. 多分辨率边界网格处理策略要求网格尺寸的初始值 r_g 与 c_g 均为 2 的正整数次幂. 在实际的图像融合中, 这一要求并不苛刻.

特征相似性原则的关键在于对充分近与充分远的判断. 对于某一边界网格 $f(i, j)$, 设

$$S_{\min} = \arg \min_{S \in \{r_domain, c_domain_A, c_domain_B\}} \|f(i, j) - v_s\| \quad (7)$$

其中, $\|\cdot\|$ 表示欧氏范数. 如果满足如下条件:

$$\frac{\|f(i, j) - v_s\|}{\|f(i, j) - v_{S_{\min}}\|} > T_{S_{\min}, S}, \forall S (S \in \{r_domain, c_domain_A, c_domain_B\} - \{S_{\min}\}) \quad (8)$$

则网格 $f(i, j)$ 与区域 S_{\min} 的距离充分近. 阈值 T_{S_1, S_2} 的计算方法如下:

$$T_{S_1, S_2} = \frac{\|v_{S_2} - v_{S_1}\|}{d_{S_1}}, (S_1, S_2 \in \{r_domain, c_domain_A, c_domain_B\}) \wedge (S_1 \cap S_2 = \emptyset) \quad (9)$$

其中,

$$d_s = \frac{\max_{f(i, j) \in S, e_{ij}=0} \|f(i, j) - v_s\| + \min_{f(i, j) \in S, e_{ij}=0} \|f(i, j) - v_s\|}{2}, S \in \{r_domain, c_domain_A, c_domain_B\} \quad (10)$$

本质上讲, 边界网格处理过程就是互补信息区域与冗余信息区域初始划分的调整过程, 经这一过程, 互补信息区域的形状将逼近图像中目标的真实形状, 网格尺寸初值对区域形状的限制被消除, 区域边界效应得以抑制. 多分辨率边界网格处理策略既保证了网格分配的准确性, 又将分配过程的计算复杂度维持在较低的水平.

基于以上原理, 本文提出如下区域鉴别算法:

Algorithm Comp_Redund_Discrimination.

Input: f, q ;

Output: $r_domain, c_domain_A, c_domain_B$.

{以 q 为参数使用 SVC 对 f 中所有网格特征向量聚类, 共计算得到 $counter$ 个聚类, 第 i 个聚类的均值向量为 $average[i]$, $1 \leq i \leq counter$;

/*互补信息区域与冗余信息区域的初始划分策略*/

计算 f 中欧氏范数最小的 n 维特征向量 f_{\min} , 其所属聚类的编号为 $i_{f_{\min}}$, $T = \|average[i_{f_{\min}}]\|$;

for ($i=1; i \leq counter; i++$)

```

    {if ( $\|average[i]\| \leq T$ )          将第  $i$  个聚类加入  $r\_domain$ ;
     else if ( $op(average[i]) > 0$ )  将第  $i$  个聚类加入  $c\_domain\_A$ ;
     else if ( $op(average[i]) < 0$ )  将第  $i$  个聚类加入  $c\_domain\_B$ ;
     else                            将第  $i$  个聚类加入未确定区域  $u\_domain$ ;
    }

```

/*函数 $op(a)$ 的功能是: 如果向量 a 的所有分量均大于 0, 则 op 返回 1; 如果向量 a 的所有分量均小于 0,

```

    则  $op$  返回-1.其他情况  $op$  返回 0*/
while ( $u\_domain \neq \emptyset$ )
    {从  $u\_domain$  中任取一个聚类  $clus$ ;
    计算  $clus$  到区域  $r\_domain, c\_domain\_A, c\_domain\_B$  的距离,此距离定义为聚类  $clus$  的均值向量对
    应的  $n$  维空间的数据点到相应区域中所有聚类均值向量对应数据点的欧氏距离的最小值;
    将  $clus$  加入与之距离最近的区域;
    }
将区域  $r\_domain, c\_domain\_A, c\_domain\_B$  中的元素由聚类变换为网格,即:提取某个区域中所有聚类
内的网格,这些网格的全体重新构成该区域,不再使用聚类的概念;
/*多分辨率边界网格处理策略*/
使用 4-邻域搜索计算边界矩阵  $E$ ;
由式(6)计算互补信息区域与冗余信息区域的特征向量  $v_s$ ;由式(9)、式(10)计算特征相似性阈值  $T_{s_1, s_2}$ ;
设置尺度因子  $scale=1$ ;
if ( $r_g \neq 1$ )    { $r_g=r_g/2; scale=2 \times scale$ ;}
if ( $c_g \neq 1$ )    { $c_g=c_g/2; scale=2 \times scale$ ;}
while ( $\neg((r_g==1) \wedge (c_g==1))$ )
    {扩展边界矩阵  $E$ ,原矩阵的每个元素复制  $scale$  次,  $E$  的尺寸为  $(r/r_g) \times (c/c_g)$ ;
    根据式(4)、式(5)重新计算网格特征  $f$ ;
     $v_s = v_s / scale$ ;     $T_{s_1, s_2} = T_{s_1, s_2} / scale$ ;
    对于每一个  $e_{ij}=1$  的网格  $f(i, j)$ ,根据式(7)计算  $S_{min}$ ,根据式(8)判断  $f(i, j)$  与区域  $S_{min}$  的距离是否充分
    近,如果充分近,则将  $f(i, j)$  加入  $S_{min}$ ,同时置  $e_{ij}=0$ ;
    设置尺度因子  $scale=1$ ;
    if ( $r_g \neq 1$ )    { $r_g=r_g/2; scale=2 \times scale$ ;}
    if ( $c_g \neq 1$ )    { $c_g=c_g/2; scale=2 \times scale$ ;}
    }
扩展边界矩阵  $E$ ;计算  $v_s = v_s / scale$ ;
计算网格特征  $f(i, j) = [|\mathbf{w}_1^A(i, j)| - |\mathbf{w}_1^B(i, j)|, |\mathbf{w}_2^A(i, j)| - |\mathbf{w}_2^B(i, j)|, \dots, |\mathbf{w}_n^A(i, j)| - |\mathbf{w}_n^B(i, j)|]^T$ , 其中,
 $i=1, 2, \dots, r, j=1, 2, \dots, c$ ;
计算每一个  $e_{ij}=1$  的单像素网格与区域  $r\_domain, c\_domain\_A, c\_domain\_B$  特征向量的欧氏距离,将单
像素网格分配到与之距离最近的区域中去;
}

```

通过以上算法可以得到冗余信息区域 r_domain 和互补信息区域 c_domain_A, c_domain_B . SVC 的计算复杂度较大,其中无向图邻接矩阵的计算复杂度尤其大.为了克服这一弱点,应选取较大的网格尺寸初值,这样可以大幅度减少网格特征数量,从而减少 SVC 的计算时间.较大的网格尺寸会导致明显的区域边界效应,这一问题将由多分辨率边界网格处理策略解决.由于非边界网格特征是准确的,从而互补信息区域与冗余信息区域的特征向量是准确的,依据准确的特征向量对边界网格进行多分辨率分配,分配过程将逐步逼近准确的区域边界.实验结果表明,算法 Comp_Redund_Discrimination 是有效的.区域鉴别算法的核心是 SVC,如果 SVC 不能提供准确的聚类结果,则后续所有处理均将失去意义.

根据 $r_domain, c_domain_A, c_domain_B$ 确定融合方案.如果网格属于互补信息区域,使用选择法生成融合图像的多分辨率表示,依据是:互补信息区域中源图像的清晰程度差异较大,必须将清晰网格加入融合图像,而放弃模糊网格.如果网格属于冗余信息区域,使用加权平均法生成融合图像的多分辨率表示,这是为了实现平滑过渡,抑制拼接痕迹的产生.选择法与加权平均法均按照层次进行.加权平均法在每一层所使用的加权系数将根

据源图像相应层的清晰程度矩阵计算.对融合图像的多分辨率表示进行小波逆变换即可得到融合图像 F .

基于以上融合规则,在不考虑 BSV 的条件下,SVC 的参数 q 成为融合效果的控制参数.如果 q 值较小,则大部分网格聚集到冗余信息区域,而且当 q 值足够小时,所有网格均属于冗余信息区域.这样,本应属于互补信息区域的网格被归入冗余信息区域,使用加权平均法融合,融合效果较差.如果 q 值较大,则聚类精细,聚类数目较多,每个聚类中元素个数较少,冗余信息区域缩小,本应属于冗余信息区域的网格被归入互补信息区域,使用选择法融合,这会导致相邻网格间的非平滑过渡,融合效果同样较差.因此,参数 q 只有在某一特定范围内选取,融合效果才会令人满意.

2.3 参数 q 的迭代搜索策略

根据第 2.2 节对 SVC 的参数 q 与融合效果之间关系的分析,可以采用迭代策略优化参数 q ,以获得令人满意的融合效果.迭代策略的关键是初始值与迭代步长的选取以及停止条件的定义,按照 Ben-Hur 等人在文献[8]中的分析,参数 q 的初始值 q_{init} 选为

$$q_{init} = \frac{1}{\max_{i_1, j_2=1, \dots, r/r_g, j_1, j_2=1, \dots, c/c_g} \|f(i_1, j_1) - f(i_2, j_2)\|^2} \quad (11)$$

迭代步长 Δq 的选取应考虑所拥有的计算资源.如果 Δq 较小,则计算时间较长,但搜索精细;如果 Δq 较大,则计算时间较短,但可能遗漏关键搜索点,甚至无法计算出令人满意的融合图像.一般可选 $\Delta q = 100q_{init} \sim 10000q_{init}$.

依据参数 q 与融合效果之间的关系,以 $q = q_{init}$ 为起点,每次迭代 $q = q + \Delta q$,则随着迭代过程的进行,融合效果逐渐令人满意.但是,如果迭代过程无限制进行下去,则融合效果将逐渐变差,因而迭代停止条件 S_cond 的定义至关重要.本文以冗余信息区域的计算准确性为依据定义 S_cond ,即:如果冗余信息区域的计算结果与图像中真实的冗余信息区域相符,则迭代停止.真实的冗余信息区域需要估计,一种简单的方法是计算两幅源图像的差图像,如果差图像像素的绝对值小于某一较小数值 ε ,则该像素属于冗余信息区域.统计冗余信息区域的像素个数,将其与像素总数 rc 的比值作为阈值 T_r ,如果迭代过程计算出的冗余信息区域所占比例在一个长度为 N 的时间窗口内不发生变化且小于 T_r ,则迭代停止.

需要说明的是, T_r 并不是真实的冗余信息区域所占比例的准确值,计算 T_r 的目的是提供另外一种融合效果控制参数.第 2.2 节已经指出,参数 q 是融合效果的控制参数.然而,参数 q 的物理意义并不明确,对于任意的多聚焦图像融合问题,难以直接确定合适的 q 值,使得融合效果令人满意.比较而言, T_r 有明确的物理意义. T_r 较大,则迭代过程计算出的冗余信息区域所占比例较大,可能覆盖一部分互补信息区域,导致清晰目标的一部分变得模糊,融合效果较差; T_r 较小,则迭代过程计算出的冗余信息区域所占比例较小,冗余信息区域的一部分被划分到互补信息区域,导致出现非平滑过渡问题,融合效果同样较差. T_r 与 q 对融合效果的影响是相似的, T_r 只有在某一特定范围内选取,融合效果才会令人满意.重要的是,与 q 相比, T_r 的估计是现实的.本质上讲,参数 q 的迭代搜索策略的意义在于将难以确定的融合效果控制参数转化为易于确定的融合效果控制参数.

完整的基于 SVC 的多聚焦图像融合算法如下:

1. 使用 à trous 小波变换基于网格计算图像特征 f .
2. 计算 q_{init} , 设置 Δq 与 S_cond , 令 $q = q_{init}$.
3. 如果 S_cond 满足, 迭代结束.
4. 以 f 与 q 为参数使用算法 Comp_Redund_Discrimination 计算 $r_domain, c_domain_A, c_domain_B$.
5. 对于 c_domain_A 与 c_domain_B , 使用选择法生成融合图像的多分辨率表示; 对于 r_domain , 使用加权平均法生成融合图像的多分辨率表示. 使用 à trous 小波逆变换计算融合图像.
6. 计算 r_domain 所占比例, $q = q + \Delta q$, 转 3.

上述算法每次迭代均计算出一幅融合图像,可将其反馈给用户,如果某次迭代计算出的融合图像效果令人满意,则可以通过交互的方式在 S_cond 未满足的条件下终止迭代过程.

3 实验结果及分析

使用两种方法融合源图像.方法 1 基于文献[7]的思路,用 \hat{a} trous 小波对源图像进行多分辨率分解,基于窗口计算像素的活性指标(activity level),根据文献[7]方法训练 SVM,依据 SVM 的输出将源图像中各个像素分配到清晰区域与模糊区域,进而采用选择法融合.方法 1 简记为 AASVM(\hat{a} trous+activity+SVM).方法 2 为第 2 节提出的基于 SVC 的融合方法,简记为 AGSVC(\hat{a} trous+grid+SVC).实验主要考察 AASVM 与 AGSVC 处理区域混叠问题和非平滑过渡问题的能力,分别从视觉效果与定量分析两个角度评价融合图像的质量.为了便于比较不同的融合规则对图像特征信息的利用效率,AASVM 与 AGSVC 均使用 \hat{a} trous 小波进行多分辨率分析.

3.1 关于区域混叠问题

图 2 给出实验中的参考图像与源图像.图 2(a)为用普通数码相机拍摄的参考图像(经 Photoshop 处理),尺寸为 768×1024.分别对左侧和右侧建筑物及周围相关景物进行模糊化处理得到图 2(b)、图 2(c)所示的源图像 A,B.



Fig.2 Reference image and source images
图 2 参考图像与源图像

实验参数: $Ft = \frac{1}{256} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$, $AL = \frac{1}{16} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, $r_g=c_g=128, n=2$, 其中,AL 为 AASVM 像素活性指标

计算模板.

图 3 给出了 AGSVC 计算出的互补信息区域与冗余信息区域,不同灰度表示不同区域.AASVM 使用线性核函数,训练样本数为 200,求解二次规划,支持向量数为 48.

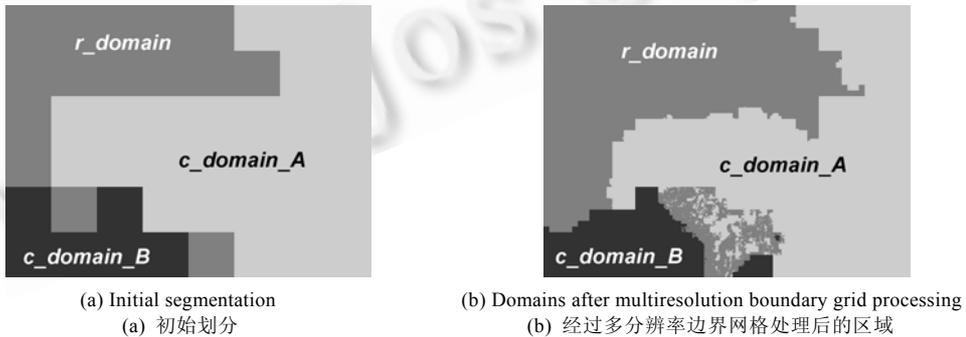


Fig.3 Complementary domain and redundant domain
图 3 互补信息区域与冗余信息区域

AGSVC,AASVM 的融合结果分别如图 4、图 5 所示.图 4(a)、图 5(a)为融合图像,图 4(b)、图 5(b)为参考图像与融合图像的差图像(128 灰度级补偿).从差图像可以看出,AASVM 的融合图像存在区域混叠问题,清晰建筑物中包含模糊建筑物的信息,而 AGSVC 的融合图像不存在此问题.区域混叠导致融合图像中出现模糊斑点,在

原始尺寸下,图5(a)中的模糊斑点清晰可见.AASVM产生区域混叠的本质原因在于清晰与模糊这两个概念本身的复杂性,AASVM所设计的样本特征并不能有效描述这种复杂性,正是这一问题导致经过训练的SVM泛化能力不足.AGSVC不涉及训练与泛化,直接对图像特征聚类,避开了概念复杂性问题.需要说明的是,AGSVC与AASVM使用同样的工具进行多分辨率分析,但这两种方法对多分辨率分析所获得信息的利用效率却不同,AGSVC挖掘语义信息的能力强于AASVM,其信息利用效率更高,融合图像质量更为理想.另外,第3.3节的定量分析将说明,AGSVC的融合图像质量是可控的.图6将图4(a)、图5(a)中右侧建筑物的黑色方框内部区域放大显示,从直观上说明了区域混叠对融合图像质量的影响.

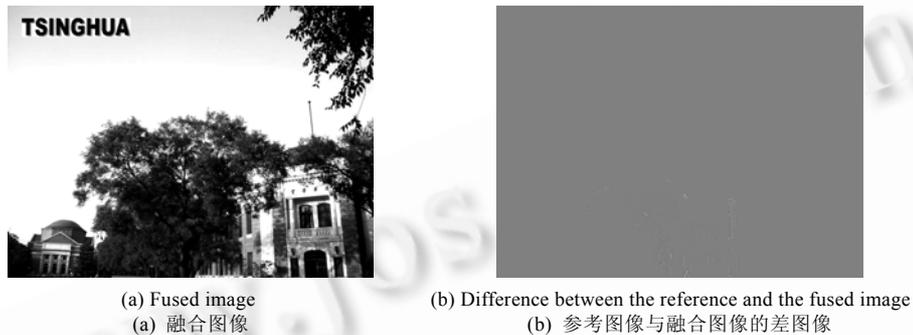


Fig.4 Fusion result computed by AGSVC

图4 AGSVC的融合结果

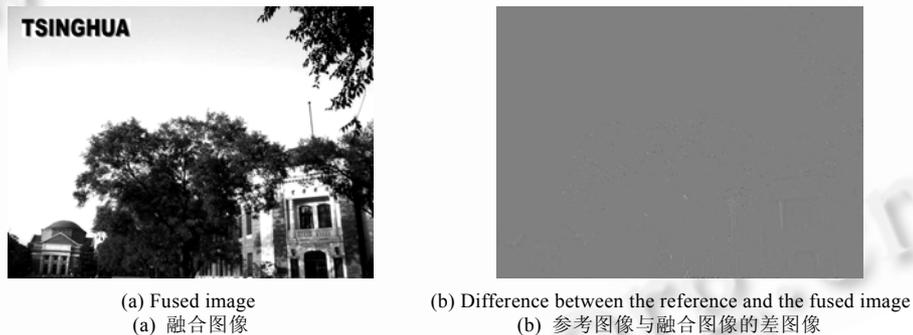


Fig.5 Fusion result computed by AASVM

图5 AASVM的融合结果

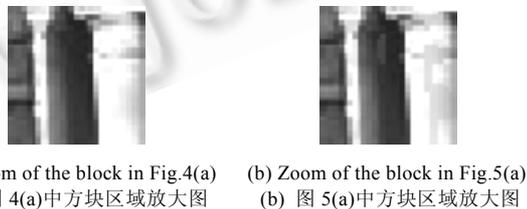


Fig.6 Demonstration of region overlapping

图6 区域混叠示意

3.2 关于非平滑过渡问题

本节通过理想图像模型的融合实验说明融合规则对目标之间过渡效果的影响.图7显示了不同融合规则所获得的不同过渡效果.图7(a)、图7(e)是由程序生成的两幅源图像,每幅源图像中条纹区域为清晰区域,均匀灰度区域为模糊区域,两区域之间过渡是平滑的.分别使用AASVM,AGSVC融合源图像,结果如图7(b)、图7(c)

所示.将两幅源图像简单平均(AVG),结果如图 7(d)所示.图 7(f)~图 7(h)分别对图 7(b)~图 7(d)中白色方框内部区域放大显示.可以看出,简单平均所得到的融合图像中区域过渡是最平滑的,然而,清晰区域对比度信息损失严重.AASVM 解决了对比度损失问题,却存在明显的拼接痕迹.AGSVC 在保证对比度信息不损失的前提下,实现了区域之间的平滑过渡.AASVM 存在非平滑过渡的本质原因在于仅将像素分为清晰与模糊两类,这等价于将源图像中的清晰目标简单拼接在一起.AGSVC 定义了互补信息区域与冗余信息区域,通过准确定位这两个区域,将需要加权平均与需要选择的像素分开,避免了在过渡区域产生拼接痕迹.

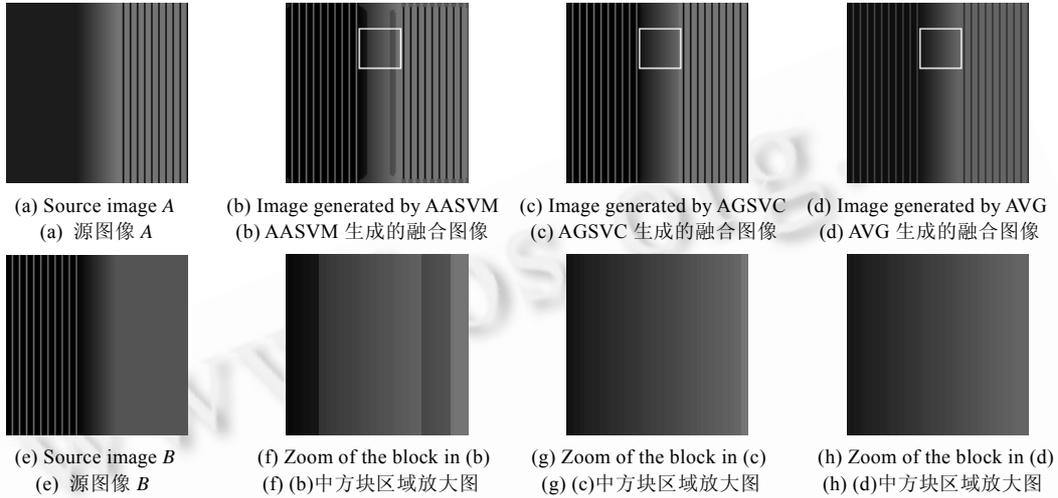


Fig.7 Examples employed to demonstrate abrupt or gradual transition brought about by different fusion rules

图 7 使用不同融合规则所获得的平滑与非平滑过渡效果示例

3.3 q 与RMSE的关系

为了定量研究 SVC 的参数 q 与融合效果之间的关系,如下定义均方根误差:

$$RMSE = \sqrt{\frac{1}{rc} \sum_{i=1}^r \sum_{j=1}^c (R(i, j) - F(i, j))^2}$$

RMSE越小,则融合图像 F 与参考图像 R 之间的差异越小,融合效果越理想.在第 3.1 节的实验中,AASVM的 $RMSE=4.61$,AGSVC在 $q=2.49 \times 10^{-8}$ 时 $RMSE=2.70$.图 8 显示了第 3.1 节实验中AGSVC的参数 q 在 $3.56 \times 10^{-12} \sim 1.05$ 范围内发生变化时 $RMSE$ 的变化情况.

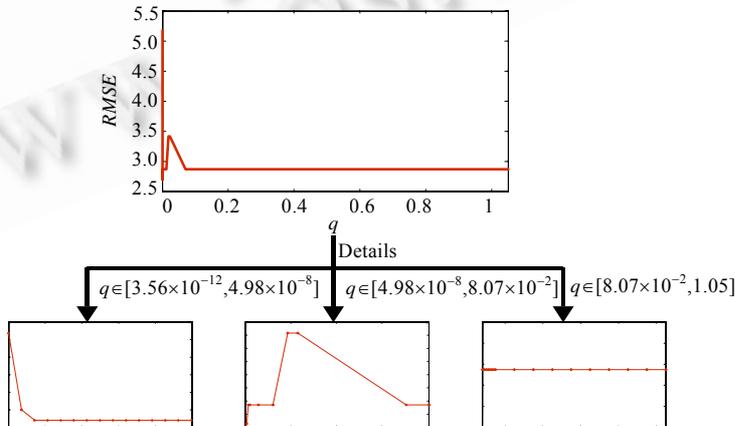


Fig.8 Relation between q and $RMSE$

图 8 参数 q 与 $RMSE$ 的关系曲线

从图 4 可以看出,使用 AGSVC 来融合多聚焦图像能够获得令人满意的视觉效果.由图 8 给出的 $RMSE$ 与参数 q 之间的函数关系可知,当 q 较小时, $RMSE$ 较大,融合效果较差.随着 q 的逐渐增大, $RMSE$ 逐渐减小,融合效果逐渐变得令人满意.而当 q 增大到某一点时,若 q 继续增大,则 $RMSE$ 开始增大,融合效果开始变差.在这一过程中, $RMSE$ 的减小速率大于其增大速率.虽然 $RMSE$ 在变化过程中有波动,但总的趋势是先减小至最小值然后增大至稳态值.图 8 从定量角度证实了第 2.2 节中关于 q 与融合效果之间关系的分析,也证实了第 2.3 节中迭代算法的合理性.

3.4 AASVM与AGSVC的性能比较

为检验 AGSVC 的适应性,除了进行如图 2 所示的 *Campus* 图像的融合实验以外,另外选择有代表性的 4 组图像,按照与 *Campus* 图像融合实验一致的实验方法对其进行融合.图 9 给出 4 幅参考图像.AGSVC 所涉及的主要参数见表 1.AASVM 使用 *Campus* 实验所训练的 SVM.



Fig.9 Reference images labeled *Clock*, *Balloon*, *Lab* and *Pepsi* respectively
图 9 参考图像 *Clock*,*Balloon*,*Lab* 和 *Pepsi*

Table 1 Key parameters of AGSVC in all of the experiments
表 1 实验中 AGSVC 的主要参数

Para Image	Image size (r×c)	Grid size (r _g ×c _g)	Initial q (q _{ini})	Search step (Δq)	Threshold (T _r) (%)	Window size (N)	Iteration number	Final q
<i>Campus</i>	768×1024	128×128	3.56×10 ⁻¹²	3.56×10 ⁻⁹	52.94	5	8	2.49×10 ⁻⁸
<i>Clock</i>	128×128	16×16	7.82×10 ⁻⁸	7.82×10 ⁻⁵	58.03	5	11	7.82×10 ⁻⁴
<i>Balloon</i>	480×640	32×128	1.00×10 ⁻⁹	1.00×10 ⁻⁶	65.86	5	9	8.03×10 ⁻⁶
<i>Lab</i>	240×320	16×32	7.09×10 ⁻⁹	7.09×10 ⁻⁶	58.29	5	12	7.80×10 ⁻⁵
<i>Pepsi</i>	128×128	16×16	4.87×10 ⁻⁸	4.87×10 ⁻⁴	45.88	5	6	2.44×10 ⁻³

图 10 比较了 AASVM 与 AGSVC 在 5 个融合实验中所能达到的 $RMSE$ 指标.可以看出,AGSVC 的融合结果更接近理想的图像.

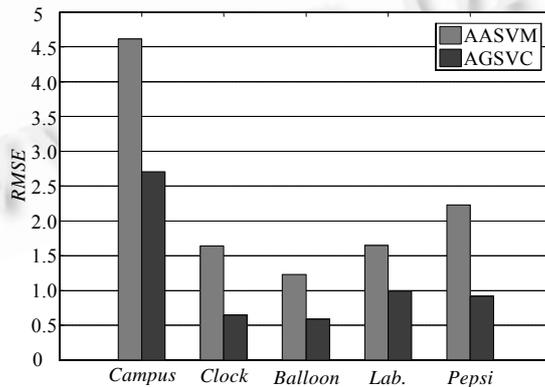


Fig.10 $RMSE$ comparison between AASVM and AGSVC
图 10 AASVM 与 AGSVC 的 $RMSE$ 比较

图 11 给出了 *Clock* 图像融合实验结果.分别对图 9 中 *Clock* 图像左侧和右侧钟表进行模糊化处理得到图 11(a)、图 11(e)所示的源图像 A、B,图 11(b)、图 11(f)分别为 AASVM、AGSVC 的融合图像,图 11(c)、图 11(g)分别

为参考图像 *Clock* 与图 11(b)、图 11(f) 的融合图像的差图像(128 灰度级补偿),图 11(d)、图 11(h)分别对图 11(b)、图 11(f) 中白色方框内部区域放大显示.从图 11(d)、图 11(h) 可以看出,AASVM 的融合图像存在区域混叠问题,*Clock* 图像右侧钟表不能被完全清晰地再现,竖状条纹存在断裂痕迹,而 AGSVC 的融合图像不存在此问题.文献[7]给出的融合图像同样存在区域混叠问题,如文献[7]的图 4(b) 右侧钟表的上边框.Li 等人近期的文章^[15] 将图像块作为 SVM 分类的基本单位,这较文献[7] 中直接对像素分类的方法更为合理.然而,文献[15] 的图像块尺寸是固定的,这不适合处理目标形状不规则条件下的图像融合问题.

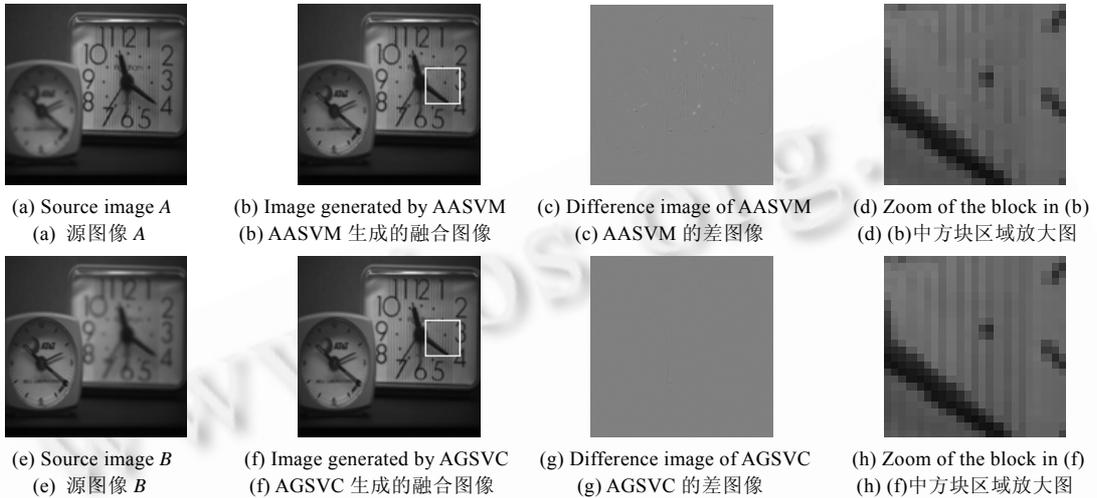


Fig.11 Images involved in the *Clock* experiment

图 11 *Clock* 实验中的相关图像

图 12 从左至右依次给出 *Clock* 实验中 AGSVC 的 11 次迭代所计算出的互补信息区域与冗余信息区域,图 12 充分说明了区域鉴别算法 *Comp_Redund_Discrimination* 对 *Clock* 图像中真实的互补信息区域与冗余信息区域的准确逼近能力.



Fig.12 Demonstration of domain evolution during the 11 times of iteration of AGSVC in the *Clock* experiment

图 12 *Clock* 实验中互补信息区域与冗余信息区域在 AGSVC 的 11 次迭代中的演化示意

4 结论与后续工作

本文以多聚焦图像融合为背景,从聚类角度研究了融合规则问题.提出了基于 SVC 的图像融合规则,解决了基于 SVM 的融合规则存在的区域混叠问题与非平滑过渡问题.理论分析及实验结果均表明,SVC 用于图像融合问题是合适的.本文从定性和定量角度研究了 SVC 的参数 q 与融合效果的评价参数 $RMSE$ 之间的关系.与基于 SVM 的融合规则相比,基于 SVC 的融合规则所能获得的融合效果是可控的.基于聚类的图像融合方法对输入源图像的依赖性较小,适用范围更广.下一步需要重点解决本文算法的时间复杂度较大的问题以及本文算法向图像融合的其他领域扩展的问题.

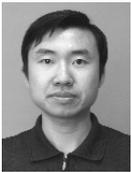
References:

- [1] Li H, Manjunath BS, Mitra SK. Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, 1995,57(3):235–245.
- [2] Burt PJ, Adelson EH. The Laplacian pyramid as a compact image code. *IEEE Trans. on Communications*, 1983,31(4):532–540.

- [3] Zhang Z, Blum RS. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. Proc. of the IEEE, 1999,87(8):1315–1326.
- [4] Petrović V. Multi-Level image fusion. In: Dasarathy BV, ed. Proc. of the SPIE 5099, Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications. 2003. 87–96. <http://spiedigitallibrary.aip.org>
- [5] Zhang ZL, Sun SH, Zheng FC. Image fusion based on median filters and SOFM neural networks: A three-step scheme. Signal Processing, 2001,81(6):1325–1330.
- [6] Li ST, Kwok JT, Wang YN. Multifocus image fusion using artificial neural networks. Pattern Recognition Letters, 2002,23(8): 985–997.
- [7] Li ST, Kwok JT, Tsang IW, Wang YN. Fusing images with different focuses using support vector machines. IEEE Trans. on Neural Networks, 2004,15(6):1555–1561.
- [8] Ben-Hur A, Horn D, Siegelmann HT, Vapnik V. Support vector clustering. Journal of Machine Learning Research, 2001,2(12): 125–137.
- [9] Vapnik V, Written; Zhang XG, Trans. The Nature of Statistical Learning Theory. Beijing: Tsinghua University Press, 2000 (in Chinese).
- [10] Horn D. Clustering via Hilbert space. Physica A, 2001,302(1-4):70–79.
- [11] Chiang JH, Hao PY. A new kernel-based fuzzy clustering approach: support vector clustering with cell growing. IEEE Trans. on Fuzzy Systems, 2003,11(4):518–527.
- [12] Lee J, Lee D. An improved cluster labeling method for support vector clustering. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005,27(3):461–464.
- [13] Núñez J, Otazu X, Fors O, Prades A, Palà V, Arbiol R. Multiresolution-Based image fusion with additive wavelet decomposition. IEEE Trans. on Geoscience and Remote Sensing, 1999,37(3):1204–1211.
- [14] González-Audicana M, Otazu X, Fors O, Seco A. Comparison between Mallat's and the 'à trous' discrete wavelet transform based algorithms for the fusion of multispectral and panchromatic images. Int'l Journal of Remote Sensing, 2005,26(3):595–614.
- [15] Li ST, Wang YN. Multifocus image fusion using spatial features and support vector machine. In: Wang J, Liao X, Yi Z, eds. LNCS 3497, Berlin: Springer-Verlag, 2005. 753–758.

附中文参考文献:

- [9] Vapnik V,著;张学工,译.统计学习理论的本质.北京:清华大学出版社,2000.



张诚成(1980—),男,黑龙江北安人,硕士,主要研究领域为机器学习,图像融合,计算机视觉。



胡金春(1972—),男,博士,副教授,主要研究领域为模式识别,智能控制,图像处理,嵌入式系统。