

舌诊信息化过程中的 质苔分离方法

楚宇燕 王博亮 王 雷

关键词 HSV 色彩空间 最佳阈值迭代算法 彩色特征组合

摘要 应用计算机分析舌像,在舌像信息辅助下推断中医证型的过程中,实现舌质、舌苔区域的自动化分离是前提条件,分离结果直接影响到后续舌质、舌苔颜色分类识别的质量,最终影响证型推断的准确性。在质苔分离实验中采用两种方法,即基于 HSV 色彩空间的最佳阈值迭代算法和基于彩色特征组合的分割方法,根据两者各自特点,综合叠加运用这两种方法,有效的解决质苔分离问题。

The Method of Separation of Tongue Coating and Tongue Texture in the Process of Tongue Diagnosis Informatization / Yu-yan CHU, Bo-liang WANG, Lei WANG // China Digital Medicine. -2007, 2 (9): 29 to 33

Keywords HSV color space, optimal threshold iterated algorithm, color characteristics

Abstract The separation of the coating and the texture of the tongue is crucial and prerequisite for the TCM-diagnosing from the information of the tongue which derived from the analysis of the tongue picture with the computer. The quality of the separation will directly affect the classification and recognition of the color of the tongue coating and tongue texture in the following step, at last, the accuracy of syndrome differentiation. Hence, this article concentrates on analysis of the automatic separation of coating and the texture of the tongue. Two methods are comprehensive applied in the separation, first is Optimal Threshold Iterated Algorithm based on HSV color space, and second is based on the color characteristics. Through comprehensive use of the two methods according to their own characteristics, we achieve the purpose of the perfect separation.

Corresponding author Department of Computer Science, Xiamen University, Xiamen 361005, Fujian Province, P.R.C.

Project supported by National Natural Scientific Fund(No: 60601025), MOH-Fujian Prov. United Fund(No: WKJ2005-2-001) and Scientific Planning of Xiamen Municipal

1 引言

舌质舌苔自动化分离是舌诊客观化研究过程中重要步骤之一,也是后期舌质和舌苔颜色识别的重要前提。目前国内许多学者也在这方面做了卓有成效的研究,其中主要有:上海中医药大学孙扬提出的基于分裂-合并方法^[1]的中医舌像区域分割

基金项目:国家自然科学基金(编号:60601025);卫生部联合基金(编号:WKJ2005-2-001);厦门市科技计划(编号:3502220055003)
厦门大学计算机科学系,361005,厦门市思明南路422号

算法, 该算法对于苔厚的舌像具有较好的分割效果, 但是对于舌苔与舌质交错的苔薄舌像, 分割的误差较大; 也有诸如模糊C均值(FCM)聚类算法^[2]、K-means 分层聚类^[3]等方法被用于质苔分离工作。这些算法对苔厚舌像的质苔分离的效果都不错, 但是彩色聚类是基于灰度阈值分割的多维扩展, 通常对每一颜色分量上的直方图分别采用阈值分割, 再将结果组合起来。由于彩色直方图阈值难于确定以及色彩映射导致信息耗散, 采用这种方法通常容易出现聚类相互重叠, 尤其对于苔色与舌色交错, 舌苔分布不均匀、不连续或多色彩的舌像, 难以确定聚类中心, 造成错分, 乱分的现象; 浙江大学陈海燕提出的基于多色彩通道动态阈值的分离算法^[4], 采用多色彩通道择优动态确定阈值的双峰分离法, 而择优只是相对其他色彩通道来说最优, 由于舌像颜色组合繁多, 有时是不能仅靠选择一个最优色彩通道就能解决质苔分离问题的。本文综合利用颜色信息, 提出基于HSV色彩空间的最佳阈值迭代算法结合彩色特征组合的质苔分离方法。

2 基本原理及实现

本文从分析基于HSV色彩空间的最佳阈值迭代算法和基于彩色特征组合的分割方法着手, 得出它们各自的特点, 然后加以综合叠加运用。在进行质苔分离时先引入了HSV色彩空间, 考虑到HSV色彩空间中H分量代表颜色类别, S分量代表颜色深浅程度的特点, 于是在H和S分量上叠加应用最佳阈值迭代算法, 从而可以达到把舌质舌苔初步分离的目的。在初步分离得到的舌质部分中运用基于彩色特征组合的分割方法, 分离出更精

确的最终舌质部分, 然后对舌像图片(本文中提到的舌像图片均为舌体分割良好, 并作空间变换之后的舌像图片)运用代数运算得到最终舌苔部分, 从而完成质苔分离工作。

2.1 基于HSV色彩空间的最佳阈值迭代算法

2.1.1 HSV色彩空间 HSV色彩模型^[5]

是一种更适合于肉眼识别的模型。在HSV色彩空间, H分量是色调, 色调决定了彩色光的光谱成分, 反映了彩色光在“质”方面的特征; S分量是饱和度, 饱和度是某种彩色光纯度的反映, 纯光谱色的含量越多, 其饱和度也就越高; V分量是明亮度, 明亮度决定了彩色光的强度, 是彩色光在“量”方面的特征。

在HSV色彩模型中, 如图2所示, 色调(H)处于平行于顶面的色平面上, 它们围绕中心轴V旋转和变化, 红、黄、绿、青、蓝、品红六个标准色分别相隔60度。色彩明亮度(V)沿中心轴V从下至上变化, 中心轴底端呈黑色(V=0), 顶端呈白色(V=1), 它们表示无彩色系的灰度颜色。色彩饱和度(S)沿水平方向变化, 越接近中心轴的色彩, 其饱和度越低, 正中心的色彩饱和度为零(S=0), 与最高明度的V=1相重合, 最高饱和度的颜色则处于最外边缘线上(S=1)。

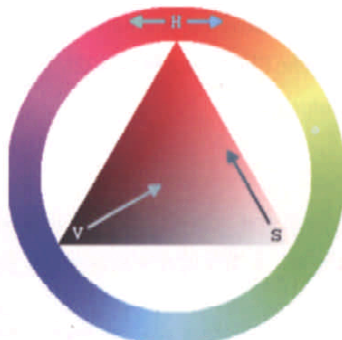


图1 HSV色彩轮

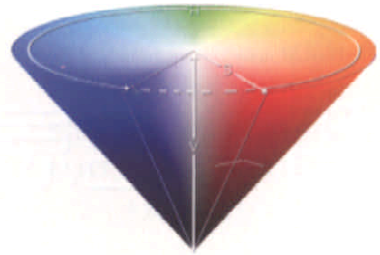


图2 圆锥形 HSV 色彩模型

彩色图像中的每个像素由红(R)、绿(G)、蓝(B)三基色按一定比例合成而表示。通过线性或非线性变换可以从三基色计算出色调(hue), 饱和度(saturation), 明亮度(value)等各种彩色特征。设RGB空间的颜色值的每个分量范围为[0,1...255], RGB色彩空间转化成HSV色彩空间如下:

$$H = \begin{cases} \arccos \frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} & B \leq G \\ 2\pi - \arccos \frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}} & B > G \end{cases}$$

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R+G+B)}$$

$$V = \frac{\max(R,G,B)}{255}$$

2.1.2 最佳阈值迭代算法

在很多情况下, 噪声的干扰使阈值的选取位置难以判定或者结果不稳定。本文采用最佳阈值迭代算法^[6], 有效减少或消除了噪声对阈值g的影响。

设有一幅混入噪声的图像g(x,y)是由原始图像f(x,y)和e(x,y)叠加而成的, 即:

$$g(x,y) = f(x,y) + e(x,y)$$

这里假设各点的噪声是互不相关的, 且均值为零, 标准差为E。通过阈值分割将图像分为目标g₁和背景g₂两部分。由于噪声是随机作用于图像的像素点上, 则可以认为在g₁和g₂图像上噪声干扰仍为e(x,y), 即:

$$g_1(x,y) = f_1(x,y) + e(x,y)$$

$$g_2(x,y) = f_2(x,y) + e(x,y)$$

在迭代算法中, 需要对分割出的图像

分别求其灰度均值,则:

$$E\{g_1(x,y)\} = E\{f_1(x,y) + e(x,y)\} = E\{f_1(x,y)\}$$

$$E\{g_2(x,y)\} = E\{f_2(x,y) + e(x,y)\} = E\{f_2(x,y)\}$$

上式说明随着迭代次数的增加,平均灰度值将趋向于真值。用这种算法求得的最佳阈值不受噪声干扰的影响,不用人工干预,自动分析完成。阈值的自动改进策略是此算法的关键。算法步骤如下。

(1) 选择一个初始阈值的估算值 $T^0 = \{T^k | k=0\}$

$$T^0 = \frac{Z_{\min} + Z_{\max}}{2}$$

式中, Z_{\min} , Z_{\max} 分别表示图像中的最小和最大灰度值。

(2) 利用阈值 T^k 把图像分成两部分, 分别是 R_1 和 R_2 , 其中

$$R_1 = \{f(x,y) | f(x,y) \geq T^k\}$$

$$R_2 = \{f(x,y) | 0 < f(x,y) < T^k\}$$

(3) 计算区域 R_1 和 R_2 的灰度均值 Z_1 和 Z_2 , 其中

$$Z_1 = \frac{\sum_{f(i,j) \geq T^k} f(i,j) \times N(i,j)}{\sum_{f(i,j) \geq T^k} N(i,j)}$$

$$Z_2 = \frac{\sum_{f(i,j) < T^k} f(i,j) \times N(i,j)}{\sum_{f(i,j) < T^k} N(i,j)}$$

式中, $f(i,j)$ 是图像上 (i,j) 点的灰度值, $N(i,j)$ 是 (i,j) 点的权重系数, 一般 $N(i,j) = 1.0$ 。

(4) 选择新的阈值 T^{k+1} 。

$$T^{k+1} = \frac{Z_1 + Z_2}{2}$$

(5) 如果 $T^k = T^{k+1}$, 则结束, 否则 $k = k + 1$, 转步骤(2)。

2.1.3 基于 HSV 色彩空间的最佳阈值迭代算法的质苔分离 由于拍摄时的角度不同, 光线变化等原因, 使得舌像图片中可能会存在光线直接反射的亮

点和完全不反射的暗点, 这些都是信息错误的点。结合对大量舌像颜色信息的统计分析, 发现这些点都有比较明显的门限值, 本文采取门限值的方法暂时把亮点和暗点都用背景色代替, 这样做不但不影响质苔分离的进行反而有助于质苔的有效分离。

对于一幅舌像图片, 引入 HSV 色彩空间, 在 H 分量上应用最佳阈值迭代算法, 称为一次分离; 再在一次分离后的舌质部分图片 S 分量上应用最佳阈值迭代算法, 称为二次分离。以一幅黄白苔淡红舌舌像为例, 一次分离和二次分离后舌质部分图片如下:

图 3(a) 是一幅周边显出淡红舌质, 舌中后部显褐黄苔, 褐黄苔周围又显白厚苔的舌像。从图 1 的 HSV 色彩轮可以看出 H 即红、黄、绿、青、蓝、紫等分量, 所以首先在 H 分量上应用最佳阈值迭代算法可以先把黄苔, 褐苔等从舌像中分离出来, 即一次分离, 如图 3(b) 所示; 又因为 S 是指某种色彩含有该色彩份量的多少的程度, 所以一次分离后, 继续在 S 分量上应用最佳阈值迭代算法, 灰度值大于最佳阈值即为舌质部分, 反之为舌苔部分, 以此来区分舌质和白苔, 即二次分离, 如图 3(c) 所示。到目前为止, 已实现了在 HSV 色彩空间对舌质舌苔的初步分离工作。

2.1.4 基于 HSV 色彩空间的最佳阈值

迭代算法质苔分离的结果分析 上述图 3 中分步显示了在 HSV 色彩空间的 H, S 颜色分量上叠加运用最佳阈值迭代算法所得的舌质部分图像, 可以看出初步分离效果良好。说明基于迭代, 逼近而产生的最佳阈值能区分出舌质和舌苔的主要区域所在, 但是在图像的细微处还没有很好的区分度。接下来, 本文应用基于彩色特征组合的分割方法对初步结果进行优化, 从而使质苔分离结果趋于完善。

2.2 基于彩色特征组合的彩色图像分割方法

2.2.1 彩色特征集 彩色图像通常用 R, G, B 分量值来表示。但是 R, G, B 这 3 个分量之间常有密切相关性, 直接利用单一分量常常不能得到所需的结果。以图 3(a) 黄白苔淡红舌的舌像的 G 分量为例, 在 G 分量上淡红舌质和白苔的反差对比强烈, 而淡红舌质和黄苔的反差对比就不明显, 所以单靠 G 分量很难较好的分离舌质和舌苔。同样道理, 单一的 R 分量和 B 分量也存在类似问题。但是色彩信息又是分析彩色图像的一个很重要的出发点, 怎样综合利用便是关键所在。

Ohta 等人通过对区域分割的系统研究, 提出用动态 K-L 变换法把一幅彩色图像三刺激值 R, G, B 变成一组相互正交的的彩色特征集 $\{l_1, l_2, l_3\}$ ^[6]。对于彩色图像不同的分割目的, 可以选

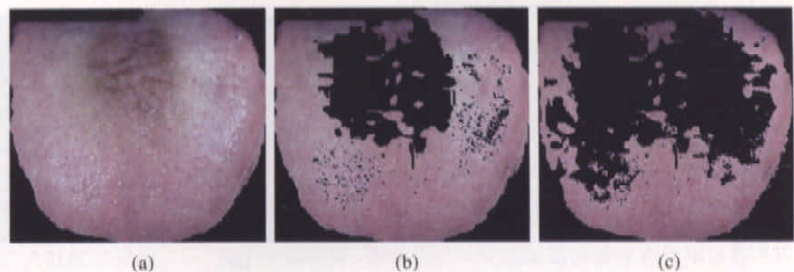


图 3 运用第一种方法进行质苔分离的效果
(a) 黄白苔淡红舌舌像图片; (b) 一次分离后; (c) 二次分离后

用不同的彩色特征组合。正交彩色特征集 $\{I_1, I_2, I_3\}$ 如下:

$$\begin{cases} I_1 = (R + G + B)/3 \\ I_2 = (R - B)/2 \\ I_3 = (2G - R - B)/4 \end{cases}$$

2.2.2 基于彩色特征组合的彩色图像分割方法用于质苔分离 由于舌质、舌苔本身的颜色特征比较多,所以二者搭配组合的舌像的颜色特征更为复杂。单单运用 HSV 色彩空间的最佳阈值迭代算法对舌像进行质苔分离,分离效果还不是很理想。在此,本文借鉴彩色图像分割的一些方法,应用彩色特征组合对舌像进行更进一步的质苔分离,称为三次分离。通过对不同舌像的 R, G, B 颜色分量相关性的分析研究,本文提出了用于进一步判断彩色舌像质苔分离分界点的公式,如下所示:

$$\begin{cases} I_2 > 0 \\ I_3 < I_4/2 \\ I_4 < 0 \end{cases}$$

其中 $I_1 = (R + B)/2$ $I_2 = (2G - R - B)/4$ $I_3 = G - B$

以一幅灰白苔淡红舌舌像为例,二次分离和三次分离后舌质部分图片见图 4。

2.2.3 基于彩色特征组合的彩色图像分割方法用于质苔分离的结果分析 图 4 (a)是淡红舌质,舌体的中后部有灰白苔的舌像图片,图 4 (b)是运用基于 HSV 颜色空间的最佳阈值迭代算法之后的舌像图片,可以看出此时的效果还不理想,在舌体后部还有部分舌苔没有和舌质部分区分开。再对图 4 (b)进一步应用基于彩色特征组合的彩色图像分割方法进行质苔分离,结果如图 4 (c)所示,得到了良好的舌质部分图片。

对于图 3 (a)进行三次分离,所得图像效果和图 3(c)的图像效果相差无

几。但是对于图 4 (a),进行三次分离后,所得图像效果远好于二次分离后的图像效果。说明本文提出的彩色特征组合公式很好的表达出了舌质部分的颜色特征,尽量使其与舌苔部分区分开。在基于 HSV 色彩空间的最佳阈值迭代算法基础上,再综合叠加运用基于彩色特征组合的分割方法,使得质苔分离效果趋于完善。

3 实验结果及对比

本文给出的基于彩色特征组合的公式可以很好的表达出舌质区域颜色,但是舌像颜色变化诸多,一些被医生认为是本属于舌苔的颜色块也有可能被本文提出的彩色特征组合公式表达出来。基于 HSV 颜色空间的最佳阈值迭代算法也同样可能存在一些本属于舌苔的区域被并入舌质部分的问题。不过这些现象只是对于部分舌像存在,我们综合叠加应用两种方法,进行一、二、三次分离,舌像的质苔分离效果得到很大提高,舌质和舌苔被错分,乱分的程度也降到最低。本文先应用第一种方法再应用第二种方法,使得质苔分离过程由粗到细,清晰可见。

实验中,我们以厦门大学医学院中医系获得的 1030 幅已经按舌质颜色,舌苔颜色及舌苔厚薄分类的舌像图片作为实验图源,质苔分离前,对舌像图片进行空间变换预处理,以舌

体的长、宽作为舌像图片的长、宽,此时的舌像图片大小约是 $230 \times 250 \times 24\text{bits}$ 。通过本文上述的分析介绍及实验结果表明,质苔分离后的舌质、舌苔部分和医生判断的舌质、舌苔部分最接近。

本文选取各类舌像图片共计 50 幅,采用五种质苔分离方法对其进行质苔分离,五种方法的横向比较情况如表 1 所示。

本文的质苔分离方法仅在一种情况下分离效果比较特殊,就是如果舌像图片整个就一片淡白(比如重淡白舌质,轻白苔),根本不分舌质、舌苔,用本文的方法所做的分离结果是舌苔部分图片和舌像图片几乎是一样的,舌质部分几乎不显示舌质颜色的信息。此时我们可以称此舌像图片质苔无明显分界点,但是同时又可确认此时的舌质颜色就是淡白舌,这正好是后续舌质颜色分析的一项内容,那么提前完成后续工作的一些内容,也可以说我们做质苔分离的最终目的已经达到。所以这种情况的出现并不影响本文质苔分离的质量。

4 结论

本文研究了基于 HSV 色彩空间的最佳阈值迭代算法结合彩色特征组合分割方法在质苔分离中的综合叠加应用。两者综合叠加运用,不仅对于厚

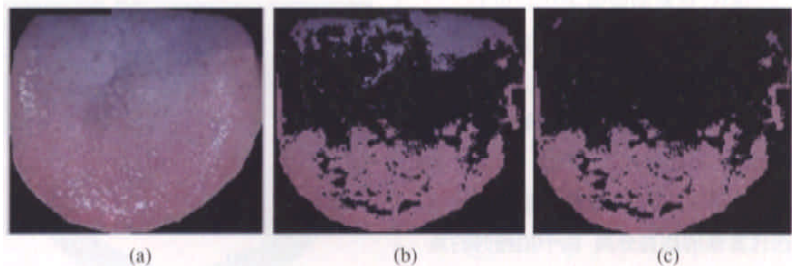


图 4 运用第二种方法对第一种方法所得的处理结果进行修正

(a) 灰白苔淡红舌舌像图片; (b) 二次分离后; (c) 三次分离后

表1 五种质苔分离方法比较

	RGB 定阈值	分裂 - 合并算法	多色彩通道动态阈值	基于L*a*b*色彩空间聚类算法	最佳阈值迭代算法结合彩色特征组合法
平均耗时	<1	<1	2	4	<1
对质苔交错的处理	弱	较弱	较强	较强	强
图像处理满意度	70%	76%	80%	90%	96%
优点	简单, 直观	区域连通性强	考虑了不同色彩空间的多个颜色分量特点	自适应能力强	运用不同颜色空间多方面颜色信息, 较全面
不足	适应性不强, 只适应某些特征明显的舌像	对质苔交错图片区分能力弱	根据面积和分布信息区分质苔, 存在不确定性; 没有综合运用多个颜色分量的不同特点	聚类中心设置影响分离结果; 对多颜色杂合舌像图片会出现错分, 乱分现象	遇到重淡白舌质, 轻白苔舌像图片, 不能明确显示质苔两部分

苔的舌像分离效果良好, 而且对于舌苔薄的舌像, 即薄而“见底”, 苔色与舌色交错, 舌苔分布不均匀、不连续的舌像分离效果也良好, 从而整体上使质苔分离效果大大提高。质苔分离的准确率直接影响着后续阶段舌质、舌苔颜色的分类识别问题, 质苔分离效果越好, 舌质、舌苔颜色的分类识别就越准确。那么本文的研究就为后续的舌像分析工作打好了基础。将来的研究方向中, 一方面是关注本文方法应用的稳定性, 是不是还存在目前未知的不足处, 以待发现问题再进一步

步解决问题; 另一方面就是以本文方法为基础, 开始着手下一阶段的舌质、舌苔颜色分类识别工作。CDM

参考文献

[1] 孙焯, 罗瑜, 周昌乐, 许家伦, 张志枫. 一种基于分裂 - 合并方法的中医舌像区域分割算法及其实现. 中国图象图形学报, 2003, 8A(12): 1395-1399
 [2] 王爱民, 沈兰荪, 赵忠旭. 监督FCM聚类算法及其在中医舌象自动分析中的应用[J]. 模式识别与人工智能, 1999, 12(4): 480-485
 [3] 朱洁华, 阮邦志, 励俊雄等. 舌诊客观化

研究的一种图象处理方法[J]. 中国生物医学工程学报. 2001, 20(2): 132-137
 [4] 陈海燕, 卜佳俊, 龚一萍, 连奕劭. 一种基于多色彩通道动态阈值的舌苔舌质分离算法. 北京生物医学工程. 2006, 25(5): 466-469
 [5] Wikipedia. the free encyclopedia: color space overview. Available at http://en.wikipedia.org/wiki/HSV_color_space
 [6] 胡小锋, 赵辉. Visual C++/MATLAB图像处理与识别实用案例精选. 北京: 人民邮电出版社, 2004

【收稿日期: 2007年7月23日】

业界观察

中国社区卫生协会成立

本刊记者陈钧报道 中国社区卫生协会7月4日在京举行成立大会。吴仪副总理为大会专门发来贺信, 并对发展社区卫生服务作了重要指示。全国人大常委会副委员长、中国社区卫生协会名誉会长韩启德, 卫生部部长陈竺, 卫生部副部长、中国社区卫生协会会长蒋作君出席会议并讲话。

韩启德对协会成立表示祝贺, 他指出: 当前, 各级党委和政府高度重视社区卫生服务工作, 社区卫生服务取得了积极的进展和显著的成效, 中国社区卫生协会的成立面临着良好的发展机遇。希望协会在政府的领导下, 按照《章程》积极主动地开展工作, 团结广大社区卫生服务工作者和社会各方力量, 充分发挥协会专家优势资源,

努力推动社区卫生服务健康有序发展。

陈竺在讲话中指出: 社区卫生服务是当前城市卫生工作的重中之重。近年来, 在党中央、国务院的正确领导下, 全国社区卫生服务发展呈现出良好态势, 取得了可喜的成绩。目前, 在部分大中城市, 已初步建立比较完善的社区卫生服务网络; 社区卫生人才建设得到进一步加强, 队伍迅速发展, 服务能力和水平得到提高; 服务功能得到完善, 公共卫生和常见病、多发病的基本医疗服务在社区得到广泛开展。社区卫生服务在有效缓解群众看病难、看病贵问题, 促进居民健康, 深化医药卫生体制改革, 推进和谐社区建设等方面正在发挥日益重要的作用, 并受到广大群众的普遍欢迎。