

# 题目：年龄估计算法技术实现

## 一. 课题来源、目的、意义、国内外基本情况

### 1.1 课题来源

本课题来源于湖北省自然科学基金项目“基于生物二值特征和稀疏表示的年龄估计问题研究（编号 2010CDB02305）”，该项目的一部分作为本课题。

### 1.2 目的

学习大型项目的分析方法、开发方法和管理模式，积累大型项目的开发经验，同时熟悉计算机中年龄估计的基本方式和应用，提高阅读论文、查阅资料、写作论文的能力。

### 1.3 意义

完成本题目的设计有助于学生总结大学四年来所学知识及所培养的能力，同时强化自我学习的能力，为将来进一步学习和工作打下良好的基础。

### 1.4 国内外基本情况

根据我的调研，Young 和 Niels<sup>[1]</sup>可能是最早提出年龄估计的人。他们早在 1994 年就提出通过人脸图像进行年龄估计。但是他们的工作相对较为简单。他们把年龄粗略地分成：小孩、年轻人和老年人三种。Hayashi 等人<sup>[2]</sup>研究了基于 Hough 变换的皱纹纹理和人脸图像肤色分析的年龄和性别识别方法。2003 年，Iga 等人用支持向量机开发一个用于估计年龄的分类器。Lanitis 等人<sup>[3]</sup>提出一种基于脸部外观的统计模型。他们比较了不同分类器，例如 KNN、MLP、SOM 的性能，并且认为机器几乎可以和人一样估计出人的年龄。Nakano 等人<sup>[4]</sup>提出利用脸部和脖子上皱纹纹理的边缘信息来进行年龄估计。Zhou 等人<sup>[5]</sup>提出用 Boosting 的方法作为回归方法进行年龄的估计，并用实验表明该方法比基于 SVMs 的方法还要好。Geng 等人<sup>[6]</sup>提出衰老模式子空间的方法，通过学习一些代表性的子空间来建模衰老模式，这种衰老模式是用一系列的个人衰老图像定义出来的。对于未知人脸图像，通过用子空间投影的方法重构其衰老模式，从而得到该人脸图像在衰老模式下的位置，由此来估计人脸图像的年龄。最近，Guo 等<sup>[7]</sup>用子空间学习方法得到衰老流形的低维嵌入，从而提取出人脸与衰老相关的特征，并设计了一个局部可调节回归方法用于学习和估计年龄。

在国内，较少人从事年龄估计的研究，文献<sup>[8]</sup>研究了基于 Boosting RBF 神经网络的人脸年龄估计方法，先用非负矩阵分解方法提取人脸特征，然后通过 RBF

神经网络确定一个人脸图像及其相符年龄之间的估计函数。为提高神经网络的泛化能力和故障诊断的准确性，利用 Boosting 方法构造一个基于神经网络的函数序列，将它们组合成一个加强的估计函数，最后进行年龄估计。文献<sup>[9]</sup>提出一种基于人工免疫识别系统的年龄估计方法，先利用 AAM 方法自动提取用于年龄估计的人脸特征，然后利用人工免疫识别系统方法，进行人脸图的年龄估计。

年龄估计是一个比较复杂的问题。这是因为，人的年龄特征在外表上很难准确地被观察出来，即使是人用眼睛也很难准确地判断出一个人的年龄。人脸的年龄特征通常表现在皮肤纹理、皮肤颜色、光亮程度和皱纹纹理等方面，然而这些因素通常与个人的遗传基因、生活习惯、性别、性格特征和工作环境等方面相关。因此，很难用一个统一的模型去定义人脸图像的年龄，通常需要通过大量样本的学习才能较好地估计出人的年龄层次。

## 二．预计达到的目标、关键理论和技术、技术指标、完成课题的方案及主要措施

### 2.1 预计达到的目标

基于 Guodong Guo 的研究成果，编写一个能够实时处理摄像头提取图像的年龄估计系统。系统能由摄像头采集到的人脸图像输入，进而输出采集样本的年龄。

### 2.2 关键理论和技术

#### 2.2.1 BIF 算法的原理

BIF 算法主要是对 S1 层和 C1 层图像分别进行处理，它们的原理如下：

**S1 层：**首先，输入一个灰度图像，通过简单的 S1 单元后被分解。该 S1 单元和 Hubel and Wiesel<sup>[10]</sup>在最初的可视皮层中发现的经典的简单单元相对应。高波函数是针对 S1 单元，该 S1 单元为皮层简单细胞可视区域提供一个很好的模型。高波函数如下：

$$G(x,y) = \exp\left(-\frac{(x^2+y^2)}{2\sigma^2}\right) \times \cos\left(\frac{2\pi}{\lambda}X\right) \quad (2.1)$$

式中：

$$X = x \cos \theta + y \sin \theta \quad (2.2)$$

$$Y = -x \sin \theta + y \cos \theta \quad (2.3)$$

式中，X,Y 是高波滤波器的带有  $\theta$  角度的旋转， $\theta$  的范围是从 0 到  $\pi$ 。高宽比是固定的， $\gamma = 0.3$ ，有效的宽  $\sigma$ ，波长  $\lambda$  以及滤波器的尺寸  $s$  都是可以变化的。

但是，在我们这里的情况下，滤波器库从  $5 \times 5$  开始而不是  $7 \times 7$  开始。因此，合并之后的大小格子的尺寸和交叠都随之相应的变化。

方向 $\theta$  以不同的间距统一的从  $0$  到 $\pi$  变化，导致结果是不同数量的总的方向，比如说 4,6,8,10,12。方向的最好的数据是有我们年龄估计问题中的数据决定的。就像 Serre, Wolf, Poggio 的工作一样<sup>[11]</sup>，S1 分离器被安排用来形成一个尺度的棱锥，在一个系列范围内变化。但是在我们的案例中，我们发现从一个较小的尺寸中开始，比如  $5 \times 5$ ，而不是  $7 \times 7$ ，能够产生更好的结果。原因是

(1) 我们年龄估计问题中的图像尺寸大小是  $60 \times 60$ ，这远小于物体分类识别中的尺度。比如，在参考文献<sup>[12]</sup>中，所以的输入图像都被降低取样到 140 像素，并且保持了高宽比。

(2) 很小的面部细节对于使细微的年龄差距特征化很重要。另外，在 Serre 的工作中<sup>[12]</sup>带的数量已经被选定并固定为 8（因此总共 16 个尺寸），但是，我们让数据来决定带的最佳数量，从 2,4,6,8 中选择。S1 的可接受的视野类型的数量是由尺寸和方向的结果来决定的。比如，它可以是 192 中类型，当通过算法选择是 16 个尺度和 12 个方向的时候，角度为  $0^\circ$ ， $15^\circ$ ， $30^\circ$  …… $165^\circ$ 。

**C1 层：**皮层复杂细胞倾向于有较大的可接受的视野。C1 单元对应着这些皮层复杂细胞。C1 单元从先前的层，以相同的方向和尺度带在 S1 单元上面合并。每一个尺度都有一对相邻的分离器尺寸。S1 单元的尺度带的索引决定了 S1 相邻单元在 C1 单元上合并的尺寸。在之前的模型中，最大运算符“MAX”被使用作为合并的分离器。Riesenhuber and Poggio<sup>[13]</sup>探讨并证明了使用非线性的运算符“MAX”而不是线性求和运算符“SUM”的好处，这里我们提出另外一种非线性的运算符，标准分离“STD”：

$$\text{std} = \sqrt{\frac{1}{N_s \times N_s} \sum_{i=1}^{N_s \times N_s} (F_i - \bar{F})^2} \quad (2.4)$$

式中揭示了 S1 单元相邻的  $N_s \times N_s$  单元中数据的变量性。这里， $F_i$  是同一个尺度带（但是不同的分离器）的相邻两个 S1 单元在索引为  $i$  时的像素最大值。

$$F_i = \max(x_i^j, x_i^{j+1}) \quad (2.5)$$

式中， $x_i^j$  和  $x_i^{j+1}$  是在位置为  $i$  时尺度分别为  $j$  和  $j+1$  的分离器的值。 $\bar{F}$  是相邻的  $N_s \times N_s$  单元中分离器值的平均值。在另一个相邻尺度（比如同一个尺度带上）上面合并的操作符“MAX”增加了 2 维转换的容错，比如在一个小数量上面的尺度变化。“MAX”操作符使用相同方向但是不同尺度的分离器合并两个待分离的图像为一个，然后 STD 运算符在合并后的相邻  $N_s \times N_s$  上面操作一次。在之前的模型中，在合并的图像上面进行的操作符是另外一个“MAX”分离。尽管第二次“MAX”能够容许更多的波动和尺寸变化，但是它不能表示数据中的变量性。在年龄估计中，数据的相邻变化性的藐视对于细微的年龄差距可能是很重要的，

比如脸部的皱纹。

### 2.2.2 BIF 算法的实现

通过对上面 BIF 算法的理论知识的学习，我们知道，其实算法的实现主要就是要完成在 C1 层和 S1 层的处理，也就是主要完成两个方面的工作。

- Gabor 特征的提取
- 对提取的 Gabor 特征的处理。

在这个过程中，对于第一个部分，我们需要了解到在 MATLAB 平台上面的特征提取算法，并实现从 MATLAB 平台到 VC++ 平台的移植，在保证数据精度的情况下，充分认识和利用 MATLAB 以及 VC++ 在处理数据方面的优势，进一步提高图像处理的效率。而对于第二部分，我们则需要了解 BIF 算法中特征的处理流程，对于求最大值，标准差，以及如何提取特定单元上的特征有相应的处理和优化。

总的来说，BIF 算法的实现流程如下：

1. 设置核心参数，包括图像宽高，Gabor 核的参数，方向和尺寸。然后对于每一个方向上面的相邻尺度，利用 Gabor 特征提取函数提取特征，并求出最大值。
2. 对上面求出的最大值进行处理。对于不同的尺度，根据子块的大小提取出相应的索引，然后根据这些索引值及其大小，分别提取出特定单元上的特征。
3. 对上面的特征进行处理，主要是根据索引取值，求标准差，依次将生成的结果保存在一个一维数组中。
4. 多次重复上面的流程，直到对每一个方向都处理完毕，并将最终的结果保存下来，即为提取的 Gabor 特征。

## 3. 技术指标

利用摄像头捕捉人脸图像，作为输入，经本系统处理后，输出相应的样本年龄。比较样本的实际年龄和估计的年龄。若相同，则说明系统完成了年龄估计的功能。否则，未完成年龄估计的功能。

## 4. 完成课题的方案及主要措施

系统的工作流程如图 4.1 所示：

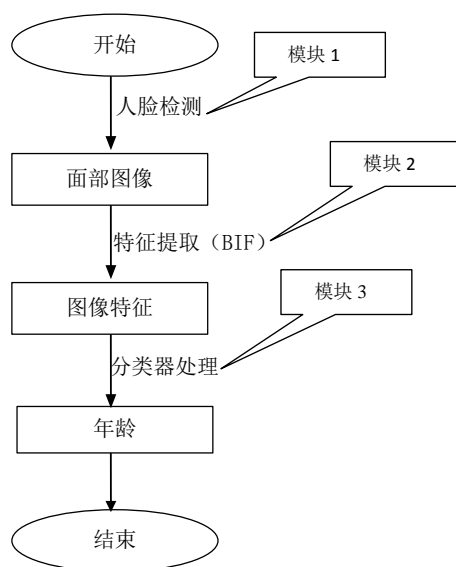


图 4.1 年龄估计系统的流程图

此流程要求我们主要完成三个模块的功能，分别是人脸检测，特征提取(BIF)以及分离器处理这三个部分，通过对每一个模块进行编码和测试，最终将各个模块的功能进行连接，完成相应的系统。下面对每一个模块都进行说明：

## 4.1 人脸检测

本原型系统使用的是一个基于 Adaboost 的人脸检测器，它能够很好的解决复杂光照、姿态变化等问题。图 4.2 给出了此人脸检测模块的框架结构。

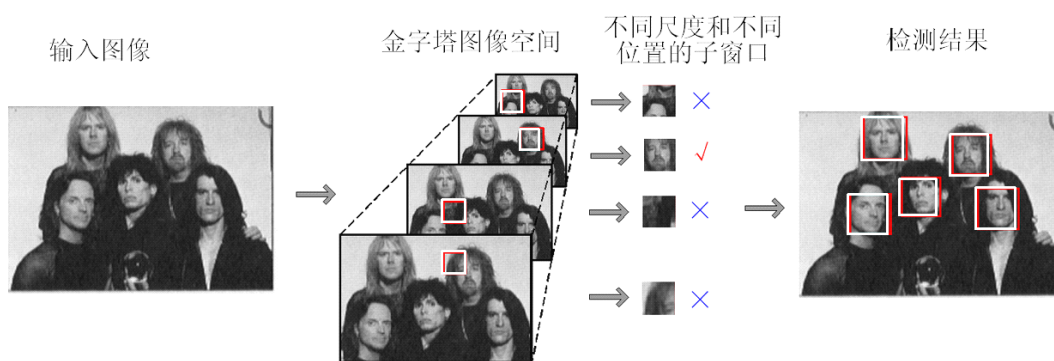


图 4.2 人脸检测算法示意图

许多的实验结果表明通过分类器的组合可以极大地开拓单个特征或者分类器的区分能力，**Boosting** 就是一种常采用的组合分类器的方式。它的一个演化算法 **AdaBoost** 已经用于人脸检测问题，并且能够在快速处理图像的同时还能达到和其他方法可比的精度。因此，本文的人脸检测系统采用了 **AdaBoost** 方法来训练分类器，通过组合一些弱分类器来形成一个最终的强分类器。

## 4.2 特征提取 (BIF 算法)

总体来说，我们这个年龄估计系统的框架如图 4.3 所示：

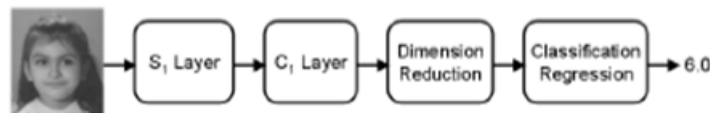


图 4.3 年龄估计的流程框架

具体来说就是：

1. 面部图像进行 S1 层和 C1 层的处理。
2. 进行维度降低用来减少维数。
3. 使用分离器或者衰减器用于学习，然后输出的结果就是估计的年龄。

在整个过程中，我们使用的 BIF 算法(Bio-inspired Features)起到了关键的作用。纵观整个处理的流程，我们会发现，BIF 算法主要针对 S1 和 C1 层进行处理，具体来说就是，首先，在 S1 层我们提取图像的 Gabor 特征，包括设定各种参数，尺寸，编写特征提取函数等等，然后提取到不同方向和尺度的特征之后，在 C1 层，我们求出相邻尺寸的最大值，然后求出它们的标准差，并将最终的结果作为提取的特征保存下来，供后面的模块使用。具体的工作流程如前面的 2.2 节中 BIF 算法的实现所示。

## 4.3 分类器训练

这里分离器训练我们主要使用 LIBSVM 来实现，LIBSVM 是一个简单、易于使用和快速有效的 SVM 模式识别与回归的软件包，它不但提供了编译好的可在 Windows 系列系统的执行文件，还提供了源代码，方便改进、修改以及在其它操作系统上应用；该软件对 SVM 所涉及的参数调节相对比较少，提供了很多的默认参数，利用这些默认参数可以解决很多问题；并提供了交互检验(Cross Validation)的功能。

LIBSVM 使用的一般步骤是：

- 1) 按照 LIBSVM 软件包所要求的格式准备数据集；
- 2) 对数据进行简单的缩放操作；
- 3) 考虑选用 RBF 核函数；
- 4) 采用交叉验证选择最佳参数 C 与 g ；
- 5) 采用最佳参数 C 与 g 对整个训练集进行训练获取支持向量机模型；
- 6) 利用获取的模型进行测试与预测。

结合以上研究内容，我设计了如下的解决问题的方案：根据系统的工作流程，

先完成这三个模块的初步的函数功能设计，对每一个模块实现的功能和算法流程有一个深刻的了解，接下来完成这三个模块的具体实现，并独立测试，测试成功后，进行系统测试，最后验证系统的年龄估计的正确率。

### 三．课题研究进展计划

综合时间，所需知识储备及工作量，制定如下的工作安排：

阶段	时间段	所要完成的工作
第一阶段	1 月 10 日~2 月 8 日	完成毕业设计的前期准备工作，包括选题，调研，阅读图像处理和模式识别教材。
第二阶段	3 月 11 日~3 月 19 日	学习论文，并翻译论文，查阅文献资料，熟悉 VC++和 MATLAB 平台。
第三阶段	3 月 20 日~3 月 28 日	主要完成系统的概要设计，包括弄清年龄估计的基本算法，工作流程中的细节，每一个模块的功能是怎样具体实现等。
第四阶段	3 月 29 日~4 月 9 日	完成第二个模块——特征提取（BIF）算法的实现，包括 MATLAB 代码的阅读和 VC++代码的编写及优化
第五阶段	4 月 10 日~4 月 30 日	完成界面部分的程序编写以及 VC++版本的 DLL 程序改写，从而实现全部的 VC 版本程序。
第六阶段	5 月 1 日~5 月 30 日	完成数据的训练，测试，并调整参数，进一步提高年龄估计的精度。
第七阶段	6 月 1 日~6 月 20 日	完成毕业论文，并完善实习日志，对整个毕业设计进行总结。

### 四．参考文献

- [1] H. K. Young, V. L. Niels, "Age Classification from Facial Images," Proceedings of IEEE Conf. on Computer Vision and Pattern Recognition. Seattle, Washington, U. S.A., 1994:762-767.
- [2] J. Hayashi, "Age and Gender Estimation Based on Wrinkle Texture and Color of Facial Images," Proceedings of the 16th International Conference. 2002:405-408.
- [3] A. Lanitis, C. Draganova, C. Christodoulou, "Comparing Different Classifiers for

Automatic Age Estimation,” IEEE Transactions on Systems, Man, and Cybernetics-partB:

[4] M. Nakano, F. Yasukata, M. Fukumi, “Age Classification from Face Images Focusing on Edge Information,” Proceedings of The 8th International Conference on Knowledge-Based Intelligent Information and Engineering Systems. Wellington, New Zealand, 2004:898–904.

[5] S. K. Zhou, B. Georgescu, X. S. Zhou, et al, “Image Based Regression Using Boosting Method,” Proceedings of the Tenth IEEE International Conference on Computer Vision. Beijing, 2005, 1:541–548.

[6] X. Geng, Z. H. Zhou, Y. Zhang, et al, “Learning from Facial Aging Patterns for Automatic Age Estimation,” Proceedings of the 14th ACM International Conference on Multimedia (ACMMM’06). Santa Barbara, CA, 2006:307–316.

[7] Guodong Guo, Yun Fu, Charles R. Dyer, and Thomas S. Huang, “Image-based human age estimation by manifold learning and locally adjusted robust regression,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 7, pp. 1178–1188, July 2008.

[8] 胡澜, 夏利民, “基于 Boosting RBF 神经网络的人脸年龄估计”, 计算机工程. 2006, 32(19):119–201.

[9] 胡澜, 夏利民, “基于人工免疫识别系统的年龄估计”, 计算机工程与应用. 2006, 26:186–188.

[10] D. Hubel and T. Wiesel. Receptive fields, binocular interaction and function architecture in the cat’s visual cortex. *Journal of Physiology*, 160:106–154, 1962.

[11] T. Serre, L. Wolf, and T. Poggio. Object recognition with features inspired by visual cortex. In Conf. on Comput. Vision and Pattern Recognit., 2005.

[12] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio. Robust object recognition with cortex-like mechanisms. IEEE Trans. Pattern Anal. Mach. Intell., 29(3):411–426, 2007.

[13] M. Riesenhuber and T. Poggio. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11):1019–1025, 1999.