## 使用大型数据库进行自动年龄估计研究

**摘要：**

在这篇论文中，我们研究一些和使用大型数据库进行人类年龄估计有关的问题。首先，我们研究了基于脸部表示的、性别在年龄估计上的影响。该面部表示将多方面的学习技术和仿生特性结合起来。其次，我们研究了使用较小的性别和年龄群而不是所有的年龄的年龄估计。在这两种情况下，我们观察了很显著的错误减少。基于这些结果，我们设计了三个框架用于高性能的自动年龄估计。不行之前的模型那样在进行年龄估计前要求人类手动的分开男性和女性，我们的工作室第一个在大型数据库上面自动的估计年龄。而且，我们使用了其中的一个框架来展示了一个数据融合的方法，该方法能够使的采用之前的模型的方法错误减少多达40%。

## 1.简介：

由于在现实世界的应用，比如电子顾客关系管理系统，安全控制和监控模拟，生物测定学，娱乐等方面的要求，人类年龄估计近来已经变成了计算机视觉方面的一个热门话题。

从计算方式来看，一个年龄估计系统通常由两个模块组成：图像表示和年龄估计。图像表示包括人体测量的模型，皱纹模型，动态外表模型，年龄子空间，从原始图像中多方面学习年龄模型，相邻二进制模式特征，和部分或者补充外表模型。给定一种表示方式，年龄估计就可以被认为是一种多级分类问题或者是一个衰减问题，或者是两个问题的综合问题。

衰老的过程和很多因素有关。比如，男性和女性有很多不同的面部衰老方式。但是到目前为止还没有研究性别在年龄估计上的影响。而且，我们想要研究在不同的年龄群上面的年龄估计性能，而不是所有的年龄上的。结果和设计一个不用用户输入性别的自动年龄估计系统有关。

我们所有的研究都要在大型数据库上面进行，因为使用小的数据库很难发现我们上面提到的这些问题。我们首先引进了在第二部分使用到的数据库，面部表示在第三部分展示，然后将研究和结果展示在第四部分。在这之后，我们在第五部分介绍一些用于自动年龄估计的框架。

## 2.数据库

我们的研究采用的是雅马哈性别年龄数据库（YGA）。它是一个最近被 采集的大型的面部数据库，包括8000张年龄从0岁到93岁、在室外的人体面部图像。这个数据库之前已经被使用作为人类年龄估计[7,31,30,11,6,10,33,12]。但是在这些所有的方法中，男性和女性都是在年龄估计之前人工手动的分开的，而这个对于一个完全的自动年龄估计系统是不实用的。

## 3.面部表示

我们探究了将仿生特征（BIF）和多方面的学习技术结合起来用于面部表示，该面部表示将在我们的研究与年龄估计有关的问题中被使用。给定了一种面部表示，支持向量机的技术会被使用用于年龄估计，比如每一个年龄都被作为一个分类的标签。基于配对的多级支持向量机之前已经成功的应用于年龄估计[10]。在这篇论文中，我们也使用支持向量机，包括各种各样的面部表示

## 3.1 仿生特征

皮层上的视觉处理被划分为一种日益复杂的、分阶级的表示。Riesenhuber

和Poggio提出了一个新的特征集合，该特征集合是从一个灵长类动物可视物体识别方式的、前馈的模型派生出来的。这个模型包括交替的简单层S和复杂层C单元的轮流交替。这个仿生特征（BIF）已经被研究用于物体分类识别[24][23][18]，和面部识别[17]。最近，仿生特征又被用于年龄估计，并在不使用预先学习的原型下表现出良好的特性[24,23,18]。这里，仿生特征在我们的研究中作为基本特征。我们发现，当使用仿生特征进行多方面的学习是，年龄估计的性能能够明显的得到提高。

## 3.2 多方面学习

这里我们研究了两种最近提出的多方面学习方法，称作费舍尔边缘分析（MFA）[32]，以及局部敏感辨别分析（LSDA）[4]，用于年龄估计的问题。之前一个较早的方法，称作正交局部保持投影[3]，已经被应用到年龄估计，在论文[7,11,10]中。在研究中，我们包含了这个方法用于比较，并且发现，当使用带有仿生特征的正交局部保持投影而不是纯粹的原始图像[7,11,10]时，能够获得一个显著的性能的提升。主成分分析（PCA）同样被使用用于比较。

## 3.3 我们的表示

我们研究了将仿生特征和多方面学习结合起来，作为年龄估计的面部表示。不像没有监控的PCA方法，这里我们研究的方法，比如LSDA和MFA，都能够在学习的时候合并标签信息。多方面学习或者子空间分析近年来市计算机视觉方面一种活跃的研究话题。通常，多方面学习方法直接应用到原始图像上[3,32,4,7,11,10,6]。但是，直接在原始图像上应用子空间分析的方法可能会有问题，因为图像的不对齐可能在学习一个可区别的、低维度的副本时造成问题。也就是说，子空间分析通常对图像的不对齐敏感。在我们的年龄估计问题中，面部图像有一个很大的年龄跨度，比如，从0岁到93岁。所以鉴于从年轻人到老人很大的面部形态变化，要对齐所有的这些图像很困难。怎样处理这个问题？使用仿生特征是一个解决办法。在提取仿生特征中，使用“MAX”操作符来从S1中获得C1特征，而这个能够容许细微的转化，旋转和尺度变化[22,23]。

我们认为，结合仿生特征和多方面学习就健壮性（健壮性比如说是对于图像的不对齐不敏感）和显著能力（显著能力指通过带有便签信息的可监控学习实现相邻的聚合）而言是有前途的。

## 4.研究和结果

基于前面部分描述的面部表示，我们研究了在年龄估计方面的两种策略:性别的在年龄估计上的影响和性别以及年龄群在年龄估计上的影响。这些策略和设计高性能的自动年龄估计方法有关，而这个会在我们接下来的研究中说明。

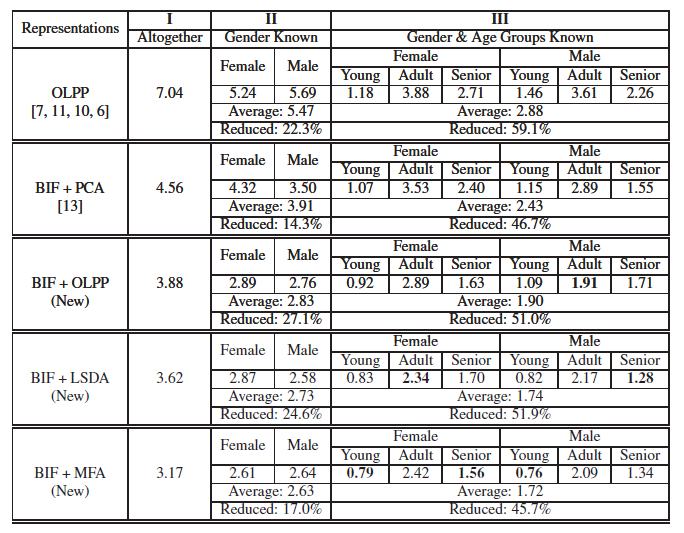
## 4.1 年龄估计中的性别未知和已知

我们首要的研究是：年龄估计的性能能够受性别多大的影响。为了研究这个，我们比较了两种情况下年龄估计的误差：1）不区分男性和女性情况下进行年龄估计。2）年龄估计分别在男性和女性上分开进行，也就是假定性别已经知道。

就我们所知，目前还没有关于性别如何影响年龄估计的研究。我们使用了各种各样的面部表示来比较在上面两种情况下的年龄估计误差。基本的表示就是使用OLPP方法从原始面部图像中学习年龄的副本，这个已经显示出年龄估计上的良好性能[7,11,10，6]。同样证实了在学习年龄副本时OLPP方法要比PCA方法更好[7]。除了像论文[7,11,10,6]中那样对男性和女性分别使用OLPP方法外，我们还不加以区分男性和女性的，对整个训练集应用OLPP方法。带有成对比较的线性支持向量机同样使用于年龄估计。年龄估计的性能通常是由平均绝对误差来衡量的[16,9]，平均绝对误差被定义为介于待估计年龄和真实年龄之间的绝对误差的平均值。就像在表1中的第1,2列所示，当运行OLPP方法于整个训练图样上时平均绝对误差为7.04，而当分别在男性和女性上应用OLPP方法时平均绝对误差分别为5.24和5.69。在男性和女性上的平均绝对误差的平均值是5.47岁，这个值是以22.3%的比率从7.04降低而来的。这个巨大的误差减少表明年龄估计受性别的影响显著。

整篇论文中，用于多方面学习和支持向量机（SVM）训练的参数是使用一个小的协调集来调整的，该协调集大概是训练数据的10%。然后多方面学习再次被执行，再在整个训练集上使用调整后的参数的SVM来再次训练。所有的结果都通过一个标准的4重交叉检查来衡量。

第二种用于年龄估计的表示是使用PCA用于维数降级的仿生特征[13]。同样，在论文[13]中只展示了分别对男性和女性进行年龄估计。当对男性和女分别进行年龄估计时，使用BIF和PCA混合的表示的平均绝对误差降低了14.3%，从平均4.56降低到3.91岁。顺便说一句，BIF和PCA混合表示要比在原始图像上的两种情况使用OLPP产生更小的年龄估计误差。



接下来，我们应用OLPP到仿生特征上。年龄估计的结果显示在表1的第三排。和PCA方法相比，OLPP方法能够在情况1下使BIF减少平均绝对误差从4.56减少到3.88。在男性和女性上（情况2）的平均绝对误差的平均值也从3.91减少到2.83岁。这证明，给定同样的仿生特征，OLPP方法比没有监控的PCA方法更好。

让我们来看看在使用“BIF+OLPP”时候情况1和情况2直接的平均绝对误差的差别。在情况2下的平均绝对误差是2.83岁，而在情况1下的平均绝对误差是3.88岁。从情况1到情况2，误差减少率为27.1%。这个显著的减少再次证明了性别在年龄估计上面的巨大影响。

接下来，我们检验了使用LSDA方法在仿生特征上进行多方面学习。在情况1下，平均绝对误差是3.62，而在情况2下，它减少到了2.73岁，减少率为24.6%。另外，“BIF+LSDA”表示比“BIF+OLPP”给出了稍微小一点的误差。

最后，我们检查了使用MFA方法在仿生特征上的多方面学习。在情况1下，平均绝对误差是3.17岁。在情况2下它减少到了2.63岁，减少率为17.0%。在这三种应用到仿生特征的多方面学习方法中，“BIF+MFA”表示有最小的误差。

综上所述，性别在年龄估计上的影响是巨大的。当在男性和女性上面分别进行年龄估计时，平均绝对误差从14.3%减少到27.1%，基于上面的五种面部表示方法。

## 4.2 性别和年龄群

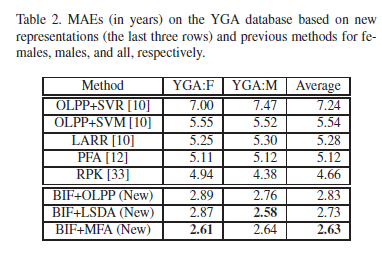
我们的第二项研究揭示了这样一个问题：如果年龄估计在一个小的性别和年龄群上进行，误差能够降低什么程度？在我们的研究中我们称之为情况3，此时性别和年龄群是已知的或者由用户提供。由在不同成长阶段的人类衰老的物理差异，我们想探究在一个年龄群而不是所有年龄上进行年龄估计时是否会获得性能的提升。从计算的角度来看，当在一个较小的年龄群上工作时，年龄估计的问题会被简化得多。

为了证实这个猜测，我们将年龄分为3个群：包括男性和女性的，青少年（0-19岁，1000张图片），成年人（20-60岁，2050张图片），老年人（61-93岁，950张图片）。然后年龄估计在每一个群上独立的进行。对于情况3下每一个群的平均绝对误差在表1中所示。所有6个群上的平均值由一个带权值的平均值来计算，考虑到群中不同的图像数量。显示在表1中的所有的结果都是通过一个标准的4重交叉检查来衡量的。

误差减少率是通过比较情况1和情况3计算出来的。给定5中表示方法，误差减少率分别是59.1%,46.7%,51.0%,51.9%,45.7%。减少时明显的，比相应的从情况1到情况2的减少要大的多。这证明，如果问题简化到工作在一个包含很小的年龄跨度时年龄估计的误差会很小。对于5中表示，情况3的结果可以被认为是在年龄估计误差上的下界。

## 4.3 和其他方法的比较

在继续接下来的研究之前，这个部分会描述一些比较。所有之前的工作[31,30,11,10,33,12]都是对男性和女性分别的进行年龄估计，类似于我们研究中的情况2。在表2中，所有的结果都是使用4重交叉检查的。后3中的结构都是基于我们新的表示。通过比较，与RPK[33]相比，新的表示方法给出了更低的误差，它的平均绝对误差的平均值为4.66岁。甚至“BIF+OLPP”方法都能够以39.3%的比率减少平均绝对误差。“BIF+MFA”使得平均绝对误差减少到2.63岁，减少率为43.6%。这些比较都说明，我们新的表示方法要比之前进行年龄估计的方法好。



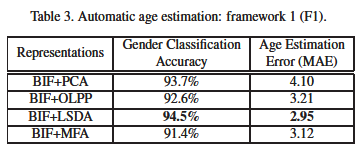
## 5.自动年龄估计

根据表1的结果，如果性别已知的话（情况2），年龄估计的性能能够得到极大的提升，而且如果性别和年龄群提供的话（情况3）能够提升的更多。事实上，但是这个信息却对于一个测试图像而言通常是未知的。为了开发一个通用的年龄估计系统，性别或者性别年龄群需要在年龄估计前被识别出来。问题是，对于完全自动的年龄估计我们能够获得怎样的性能？就我们所知，这之前还没有使用大型数据库研究过这个问题。

我们展示了若干个不同的方法来进行自动年龄估计，基于表1中给定的结果。这四种不同的面部表示是：“BIF+PCA”,“BIF+OLPP”，“BIF+LSDA”，“BIF+MFA”。在每一种方法中都使用这四种面部表示。在它们之间也进行比较来观察差异。

## 5.1 在年龄估计前预先进行性别分类

一个直接的进行自动年龄估计的方法是首先识别一个给定图像的性别，然后做和这个性别想对应的年龄估计。如果面部图像被分类为女性，那么它的年龄就是用从女性中学习的年龄估计器中得到。否则，就使用用于男性的年龄估计器。我们称这种办法为框架1（F1）。



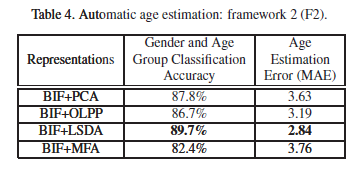
使用不同的表示的F1方法的结果在表3中显示，将表1和表3比较，可以发现：

1. 尽管性别分类不完美，但是对于所有四种表示，使用F1的自动年龄估计都能够比采用混合年龄估计获得更好的结果。
2. 使用F1进行年龄估计时最好的表示是“BIF+LSDA”，它在这四种表示中能够有最高的性别分类精度。
3. 尽管“BIF+MFA”表示的采用F1时在混合性别时有最小的平均绝对误差，但是它用于性别分类时的表现最差，因此提高了平均绝对误差。
4. 所有使用F1的表示都比性别已知时的误差大。

## 5.2 年龄估计前进行性别和年龄群的分类

鉴于我们研究中情况3下面最小的年龄估计误差，我们设计了另外一种自动年龄估计框架。它首先将一个面部图像分类到一个性别年龄群，然后在这个已经决定的群上面进行年龄估计。我们称之为框架2（F2）。如果群分类的精度足够高以至于年龄估计的误差很小（如第四部分情况3的研究所示）时，这种方法还是很合理的。F2的结果显示在表4中，从中我们可以看到：

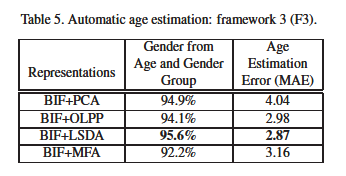
1. 使用“BIF+LSDA”表示的群分类的精度可以高达89.7%。最终的平均绝对误差可以低至2.84岁，而这个比使用F1得到的2.95岁要低。
2. 对于绝大多数表示，使用F2方式比使用F1方式的性能要好。唯一的例外情况是“BIF+MFA”。注意到，对于“BIF+MFA”群分类的精度如此之低，导致了在使用F2时它的性能很低。



## 5.3 仅仅从性别年龄群分类中使用性别

一种可以选择的使用性别和年龄群分类结果的方式是使用估计的年龄群信息来在性别上做一个新的决策。特别的，如果一个测试图像被分类为这三个群中的一个：年轻女性，成年女性，老年女性，就认定它为女性。相反，如果一个测试图像被分类为这三个群中的一个：年轻男性，成年男性，老年男性，就认定它为男性。采用这种方式的性别决定要比直接分类（比如F1）好一些吗？考虑在表5中的结果。和表3相比较，对于所有的4中表示，性别识别的精度都提高了。目前最高的精度是使用“BIF+LSDA”的表示方法，达到95.6%。甚至对于“BIF+MFA”方式，性别分类的精度也稍微从91.4%提高到了92.2%。对于大多数的表示，平均绝对误差也都减少了。唯一的例外是“BIF+MFA”。比较表5和表4，使用F3的平均绝对误差要比使用F2的“BIF+LSDA”和“BIF+PCA”表示大。F3的主要优势是它能够提高在F1中使用的、直接的两类性别分类的性别识别精度。

我们想要提到，对于年龄分类，采用RBF核心的非线性的支持向量机(SVM)要比线性的支持向量机（SVM）性能好。这个对于性别和年龄群分类也是适用的。但是对于年龄估计，线性的SVM可以和核心的SVM相比，甚至要好于核心的SVM。



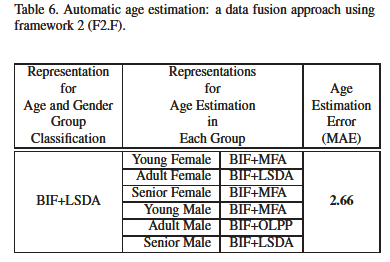
## 5.4 一个数据融合的方法

给定这三种用于自动年龄估计的框架，这四种表示中哪一种要给出最好的结果呢？表中显示，“BIF+LSDA”表示一直都是性能最好的。注意到，它对于性别和年龄群分类有最高的精度。另一方面，如果我们表1中的情况3，用于年龄估计的最小的平均绝对误差是从不同的群中不同的表示获得的。比如，“BIF+MFA”表示在这三个群中有最小的平均绝对误差：对于青年女性是0.79岁，对于老年女性是1.56岁，对于青年男性是0.76岁。而“BIF+LSDA”在这两个群中给出了最小的平均绝对误差：对于成年女性是2.34岁，对于老年男性是1.28岁。而“BIF+OLPP”则成年男性给出了最低的平均绝对误差1.91岁。因此一个自然而然的问题是：我们能否结合不同的表示来进一步提高年龄估计的性能。这个就是一个数据融合的方法。

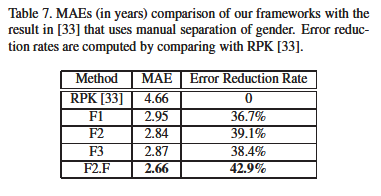
为了开发一种基于数据融合的方法，我们提出了一种简单的方法来利用到表1中情况3的结果。使用F2的融合方法能够起作用是因为仅仅F2能够充分利用6个群中的每一个群的最小的平均绝对误差。它首先将一个脸部图像分类为6个群中的一个，然后在每一个群内进行年龄估计。和5.2节的F2相比，现在在不同的群中使用不同的表示，这个由表1中的情况3所指示。只有“BIF+LSDA”表示用于性别和年龄分类。我们将这种使用F2的数据融合的方式标记为“F2.F”，并且将结果显示在表6中。平均绝对误差是2.66，这个是我们获得的最低的用于自动年龄估计的误差。

## 5.5 比较

我们采用了[33]中的方法比较了自己的自动年龄估计结果，尽管这种比较对于我们来说是不公平的，因为我们在第5节获得的所有结果都是使用自动年龄估计获得的，而之前的工作是人工手动的区分测试样本的男性和女性。这个比较的目的是为了从表3中总结出表6，并显示出改进之处。



比较的结果显示在表7中。从[33]中显示的结果是4.66岁，该数据时女性（4.94）和男性（4.38）的平均绝对误差的平均值。在我们的研究中，多种框架被使用以用于完全的自动年龄估计。在这四种表示中的最小平均绝对误差被选中作为每一个框架的代表。误差减少率很明显，从36.7%到42.9%。



## 6.结论

我们已经系统的研究了使用5中不同的面部表示和大型数据库时性别在年龄估计上的影响。我们也研究了分别在性别和年龄群上进行自动年龄估计。我们的扩展性实验表明，年龄估计受性别显著的影响，而且如果年龄估计是在包含有较小的年龄跨度和单一性别群上进行时能获得显著的误差减少。基于这些研究，我们设计了用于自动年龄估计的3种框架。一种使用框架2的数据融合方法能够给出最小的误差。和之前的其他结果相比，它能够以多余40%的比例来减少平均绝对误差。

## 参考文献

[1] Aging of the face. http://www.face-and-emotion.com/dataface/facets/aging.jsp.

[2] Electronic Customer Relationship Management (ECRM). http://en.wikipedia.org/wiki/ECRM.

[3] D. Cai, X. He, J. Han, and H. Zhang. Orthogonal laplacianfaces for face recognition.IEEE Trans. on Image Processing, 15:3608–3614, 2006.

[4] D. Cai, X. He, K. Zhou, J. Han, and H. Bao. Locality sensitive discriminant

analysis. In Proc. Int. Joint Conf. on Artificial Intell., 2007.

[5] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. In European Conf. on Computer Vision, pages 484–498, 1998.

[6] Y. Fu and T. S. Huang. Human age estimation with regression on discriminative aging manifold. IEEE Trans. on Multimedia, 10(4):578–584, 2008.

[7] Y. Fu, Y. Xu, and T. S. Huang. Estimating human ages by manifold analysis of face pictures and regression on aging features. In IEEE Conf. on Multimedia and Expo, pages 1383–1386, 2007.

[8] X. Geng, Z.-H. Zhou, and K. Smith-Miles. Automatic age estimation based on facial aging patterns. IEEE Trans. on PAMI, 29(12):2234–2240, 2007.

[9] X. Geng, Z.-H. Zhou, Y. Zhang, G. Li, and H. Dai. Learning from facial aging patterns for automatic age estimation. In ACM Conf. on Multimedia, pages 307–316, 2006.

[10] G. Guo, Y. Fu, C. Dyer, and T. Huang. Image-based human age estimation by manifold learning and locally adjusted robust regression. IEEE Trans. Image Proc., 17(7):1178–1188, 2008.

[11] G. Guo, Y. Fu, T. S. Huang, and C. Dyer. Locally adjusted robust regression for human age estimation. In IEEE Workshop on Applications of Computer Vision, 2008.

[12] G. Guo, Y. Fu, T. S. Huang, and C. Dyer. A probabilistic fusion approach to human age prediction. In IEEE CVPR-SLAM workshop, 2008.

[13] G. Guo, G. Mu, Y. Fu, and T. S. Huang. Human age estimation using bioinspired features. In IEEE CVPR, 2009.

[14] J. Hayashi, M. Yasumoto, H. Ito, and H. Koshimizu. A method for estimating and modeling age and gender using facial image processing. In Seventh Int. Conf. on Virtual Systems and Multimedia, pages 439–448, 2001.

[15] Y. Kwon and N. Lobo. Age classification from facial images. Computer Vision and Image Understanding, 74(1):1–21, 1999.

[16] A. Lanitis, C. Draganova, and C. Christodoulou. Comparing different classifiers for automatic age estimation. IEEE Trans. on SMC-B, 34(1):621–628, 2004.

[17] E. Meyers and L.Wolf. Using biologically inspired features for face processing. Int. J. Comput. Vis., 76:93–104, 2008.

[18] J. Mutch and D. Lowe. Object class recognition and localization using sparse features with limited receptive fields. In Conf. on Comput. Vision and Pattern Recognit., pages 11–18, 2006.

[19] E. Patterson, A. Sethuram, M. Albert, K. Ricanek, and M. King. Aspects of age variation in facial morphology affecting biometrics. In IEEE Conf. on Biometrics: Theory, Applications, and Systems, 2007.

[20] N. Ramanathan and R. Chellappa. Face verification across age progression. IEEE Trans. on Image Processing, 15(11):3349–3361, 2006.

[21] K. Ricanek and E. Boone. The effect of normal adult aging on standard pca face recognition accuracy rates. In International Joint Conf. on Neural Networks, pages 2018–2023, 2005.

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

[23] 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.

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

[25] J. Suo, T.Wu, S. Zhu, S. Shan, X. Chen, andW. Gao. Design sparse features for age estimation using hierarchical face model. In IEEE Int’l Conf. on Automatic Face and Gesture Recognition, 2008.

[26] K. Ueki, T. Hayashida, and T. Kobayashi. Subspace-based age-group classification using facial images under various lighting conditions. In IEEE conf. on FGR, 2006.

[27] V. N. Vapnik. Statistical Learning Theory. John Wiley, New York, 1998.

[28] A. R. Webb. Statistical Pattern Recognition, 2nd Edition. John Wiley, 2002.

[29] S. Yan, M. Liu, and T. Huang. Extracting age information from local spatially flexible patches. In IEEE conf. on ICASSP, pages 737–740, 2008.

[30] S. Yan, H. Wang, T. S. Huang, and X. Tang. Ranking with uncertain labels. In IEEE conf. on Multimedia and Expo, pages 96–99, 2007.

[31] S. Yan, H. Wang, X. Tang, and T. Huang. Learning auto-structured regressor from uncertain nonnegative labels. In IEEE conf. on ICCV, 2007.

[32] S. Yan, D. Xu, B. Zhang, H. Zhang, Q. Yang, and S. Lin. Graph embedding and extensions: A general framework for dimensionality reduction. IEEE Trans. Pattern Anal. Mach. Intell., 29:40–51, 2007.

[33] S. Yan, X. Zhou, M. Liu, M. Hasegawa-Johnson, and T. Huang. Regression from patch-kernel. In IEEE conf. on CVPR, 2008.

[34] Z. Yang and H. Ai. Demographic classification with local binary patterns. In Int. Conf. on Biometrics, pages 464–473, 2007.

[35] S. Zhou, B. Georgescu, X. Zhou, and D. Comaniciu. Image based regression using boosting method. In IEEE conf. on ICCV, pages 541–548, 2005.