融合人脸识别算法和人类

Alice J. O’TOOLE, Hervé ABDI, Fang JIANG, and, P. Jonathon PHILLIPS\*

摘要 - 最近已经证明，最先进的人脸识别算法可以超越人脸在照度变化时匹配人脸的精度。然而，算法和人类的准确性排序并不能提供关于算法和人类是否可比较地执行任务的信息，或者算法和人类是否可以融合以提高性能。在这里，我们使用偏最小二乘（PLS）回归来融合人类和算法。在第一个实验中，我们将PLS应用于参与人脸识别大挑战（FRGC）的七种算法生成的面孔对相似性得分。 PLS产生了相似性分数的最佳加权，我们使用jackknife程序测试了通用性。使用最佳权重融合算法的相似度得分比最准确的算法产生的错误率降低了两倍。接下来，将人类主体生成的相似性分数添加到PLS分析。融合人类和算法将性能提高到接近完美的分类精度。这些结果在最大化人脸识别精度方面进行了讨论，其中包括由多种算法和人类组成的混合系统。

索引术语 - 脸部和手势识别，算法和系统的性能评估，人类信息处理

# 引言

在过去的十年中，自动人脸识别算法领域已经扩展，包括在高度控制的人脸图像上运行的简单算法，以及旨在以特征为大多数安全应用的自然条件下操作的更复杂算法。 将算法从受控环境推进到自然环境中的一个特别困难的挑战是在照明的实质变化上进行操作的问题。 照明问题造成的计算困难已经在自动人脸识别[ 1-3]和人类感知文献[4-6]。

从更实际的角度来看，在人脸识别大挑战（FRGC）中，人脸识别算法在受控和非受控照明环境中的性能最近得到了评估，该算法是美国政府资助的人脸识别算法测试，旨在促进算法开发[7 ，8]。 FRGC（2004-2006）包括学术，工业和研究实验室的竞争对手。竞争对手参与了这个项目，他们通过自愿参加了六个面部匹配实验中的一个或多个测试算法，这些实验的难度各不相同。该组实验包括受控照明面部匹配实验和更难实验，其中算法使得在不同照明条件下拍摄的图像中的面部身份匹配。由于FRGC使用标准化的评估协议和通用图像集同时测试了多种算法，因此它可以为最先进的人脸识别算法的性能提供有用的，时间锁定的外观。

通过比较FRGC的受控和非受控照明实验中的算法的性能，可以清楚地看到照明问题的困难。 在这两种情况下，算法的任务是为大量的脸部对（> 1.28亿）中的每一个决定，图像是同一个人还是不同的人。 在受控照明实验中，对照中的两幅图像的照明条件相同。 在不受控制的照明实验中，一个图像在受控照明条件下拍摄，另一个在不受控照明条件下拍摄（参见图1的示例图像对）。

二十种算法在受控照明实验中竞争，以0.001的假接受率达到.91的平均验证率。 相比之下，在不受控制的照明实验中，只有7种算法参与，在.001错误接受率下平均验证率为.41。 这些实验中参与者数量和平均表现的差异表明照明问题继续挑战人脸识别算法。

在不受控制的照明实验中，相对较差的算法性能的一个相当不同的观点来自将算法与执行类似任务的人进行比较。 在最近的研究[9]中，将人脸匹配性能与参与FRGC的不受控照明匹配实验的七种算法的性能进行比较。 我们在这里详细描述了此前的研究，因为它构成了当前工作的基础。

FRGC非受控照明实验中的算法（FRCG命名法中的实验4），在16028个目标图像和8014个探头图像的所有可能对中匹配人脸身份，在受控照明条件下拍摄目标图像，并在不受控照明条件下拍摄探头图像（请参阅 图1为样本对）。 每个算法的输出是所有可能的人脸对的相似性分数的矩阵。 对于每种算法，从相似性评分矩阵生成接收机操作特性（ROC）曲线。 使用这些ROC曲线比较了七种算法的性能[参见 9为完整的结果]。



图1来自“匹配”试验的样本一对面部图像（a）参与者通过使用范围从“1”的五点量表评价图片是同一人的可能性来回答。确信他们是同一个人 “to”）5.确定它们不一样

将人类与FRGC中的算法进行性能比较的主要困难是需要进行详尽比较的难以置信的大量面孔对比较。因此，为了比较人类与算法的性能，通过选择一组最简单和最困难的对来从矩阵中对面对进行采样[9]。在目前的工作中，我们专注于最困难的图像对。然而，在这两种情况下，采样都是在基于对齐和缩放人脸图像的主成分分析（PCA）的控制算法的帮助下完成的。使用该算法，基于相似度分数来定义简单匹配对，所述相似性分数显着大于匹配脸对的分布的均值，即相同人的高度相似图像。难以匹配的配对是那些相似度分数明显低于匹配均值的配对，即同一人的高度不相似的图像。简单和困难的不匹配对被定义为相反的。

人类受试者匹配240个样本人脸对的身份，通过评估他们的确定性来确定这些人是同一个人。 人类的反应范围从“确定两幅图像属于同一个人”到“确定两幅图像不是同一个人”等五个等级。 评分数据可用于生成人体表现的ROC曲线，这与从算法性能中推导出的ROC曲线相当。

通过提取人脸匹配实验中测试的相同人脸对的算法相似度得分进行人机比较。 这些被绘制在ROC曲线上以及人类匹配精度数据[9]。 结果清楚地表明，三种算法[10-12]在困难的人脸对上超过了人类的表现。 其中，新泽西理工学院的算法[10]和卡内基梅隆大学的算法[11]已经发表。 Viisage公司的第三种算法的详细信息仅部分可用[12]。

除了发现三种算法在困难的人脸上与人类竞争之外，除了一种算法之外，所有算法在简单人脸对上都超过了人类的表现。 综合起来，这些发现表明尽管FRGC中不受控照明实验的算法性能在绝对值上可能很差，但它仍然与人类的表现相竞争。 由于人类目前在大多数应用情况下正在执行这一任务，因此这种比较是令人感兴趣的。

此前的研究构成了当前工作的基础。 尽管相对于一组算法的人类表现的定量排名提供了一个有用的基准，但这个排名并没有提供任何洞察，算法是否以与人类相似的方式识别面部。 FRGC显示，没有任何算法在非受控照明环境中完美地进行人脸识别。 我们以前的工作对人类也有同样的结果。 如果算法和人类采用不同的方法解决人脸匹配问题，那么算法和人类的适当融合可能比单一算法或多种算法的融合产生更好的性能。 事实上，以前的工作表明，融合多种人脸识别算法可以提高单个算法的性能[参见13-15]。 但是，以前的研究没有将人类和算法的表现融合在一起

在大多数人脸识别应用中，人工操作员在场并参与决策过程。 因此，通过将人脸识别能力明确纳入决策过程来优化系统性能是很自然的。 为此，我们提出了一种融合算法和人类表现的方法。

在这项研究中，我们提出了两个问题。 首先，可以通过融合FRGC不受控照明实验的算法来提高性能？ 其次，融合人类和算法是否将性能提高到算法融合所达到的水平之上？ 面部相似性的多种算法估计的可用性，以及类似的人类相似性估计，提供了以比通常可能的更系统的方式探索这些问题的可能性。 在这里，我们调查了融合来自算法和人类的面部相似性估计以改进面部匹配性能的可能性。

融合是通过偏最小二乘法（PLS）回归进行的，该统计技术将主成分分析和多元回归的特征进行了概括和组合[16,17]。 该技术用于从一组独立变量（预测变量）预测一组因变量。 尽管在模式识别文献中知之甚少，但PLS广泛用于化学计量学，感官评估和神经影像学数据分析[参见16,18,19,21。

在目前的工作中，人脸相似性的算法和人类估计是预测因子，并且单个人脸对（即同一人或不同人）的匹配状态是因变量。 PLS回归给出了一组正交因子，有时称为潜在向量{t1 ...。 tl}，来自预测变量和因变量的协方差矩阵。 这些可以用来预测因变量，通过适当的加权预测因子。 这组权重在PLS回归文献中称为Bpls [16]。 为了融合算法，使用潜在向量中规定的权重来组合来自七种算法中的每一种算法的相似性得分，以产生面对对的匹配状态的估计。 融合人类和算法时，有八个预测因子，其中七个来自算法，一个来自平均人类数据。

这些因素的预测能力通常采用交叉验证技术进行评估，例如bootstrap或jackknife程序。 所有因素，或者只是它们的一个子集，都可以用来计算因变量的预测，这些因变量是由Bpls给出的原始预测因子的加权组合得到的。 保留因子的数量越多，“学习集”的预测越好，但一般而言，对于鲁棒预测（即，对于测试集预测），较小数目的因子是最优的。

在第一个实验中，我们将PLS应用于参与FRGC不受控照明实验的七种算法所生成的相似性分数。 我们测试了在分析中发现的最佳权重的通用性，以使用折叠刀程序来预测面部匹配状态。 在第二个实验中，我们将人工生成的相似性分数添加到算法的分数中，并测量了人类估计对融合的贡献。

# 步骤

A.刺激

面部刺激是从为FRCG研究开发的大型数据库中选择的[7,8]。 不受控制的照明探头面分辨率为2272 x 1704像素。 受控照明目标面的分辨率为1704 x 2272像素。 对于目前的分析，我们使用了人类和算法之间先前定量比较的相同困难人脸对样本集[9]。 这些样本来自128,448,392对可用的样本，其中包括407,352（0.32％）个匹配对（即同一个人的图像对和128,041,040个（99.68％）的不匹配对（即不同人的图像对）。 人类可以基于与种族或年龄相关的表面面部特征进行身份比较的可能性，研究中的所有图像都是男性和女性在二十多岁时的脸部图像，所有的图像都通过性别进行匹配。

在本研究中，只有“困难的脸对”被包括在内。 这些是使用基于对齐和缩放图像的主成分分析（PCA）的控制算法来选择的。 具体而言，从相似性分数低于匹配平均值2个标准差以下的匹配对中随机抽取困难的匹配面对（n = 60）。 从非匹配对中随机抽取非匹配面对对（n = 60），这些非匹配对的相似性得分大于不匹配均值以上的2个标准差。

前面的研究支持PCA作为人类和算法的预筛选算法的有效性[9]。 PCA算法在三次实验中可靠地预测了人类“容易”和“困难”的人脸对集合[9]。 所有7种算法在PCA筛选的简单人脸对上比在PCA筛选的困难人脸上更精确[9]。 因此，PCA可以作为有用的抽样工具，尽管它不被认为是“最先进的”

B.面相似性的人类主体判断

这个实验的人体对象数据是在一个实验中收集的，在这个实验中，实验对象查看图像对，并评估图像是同一个人或不同人的图像的可能性[9]。 为了完整起见，我们概述了该研究中使用的方法。 四十九个科目（男性25名，女性26名）每次观看120副面孔2秒钟，并按以下比例对每一副面部进行评分：1）确定图片属于同一人; 2）认为这些照片属于同一个人; 3.）不知道; 4.）认为照片不是同一个人; 5.）确定图片不是同一个人。 在120对中，一半是配对，另一半是非配对。 在比赛和不匹配的条件中包括了相等数量的男性和女性对。

对于每一对面孔，平均评分是在49位受试者中计算出来的。 这个平均值作为PLS回归中该对人脸的人类相似度得分。

C.算法对人脸相似性的判断

从每个算法的16028 x 8014相似度矩阵中提取120个困难面部对的相似性得分，其被呈现给人类实验的参与者。 这些分数用作PLS回归的算法数据。

# 结果

A.实验1 - PLS算法融合

对于120个困难面部对（七十个匹配和六十个不匹配）的七种算法的相似性分数以列方式矩阵组合。 因变量是包含匹配状态的120个元素的向量，每个面对包含匹配状态（匹配为+1，非匹配为-1）。 PLS同时应用于相似性分数和匹配状态矩阵。

我们将保留的PLS因子数量从1改为5，发现三因子解决方案是最优的。 保留三个因子表明根据协方差矩阵中解释的方差比例排序的前三个潜在向量被线性组合以指定用于组合相似性分数的权重。

利用千斤顶模拟来确定稳健的性能估计。 我们从120个面对开始，依次系统地删除每个面对，用剩余的119对面重新计算PLS。 我们测试了来自“左外”面对上的119对面的PLS解的匹配状态预测。 这产生了120个广义匹配预测测试。 我们报告的错误率是根据匹配状态错误分类的残留面部对的分数。

TABLE I

WEIGHT MATRIX FOR ALGORITHM FUSION DIFFICULT FACE PAIRS



|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | *NJIT* | *Viisage* | *CMU* | *A* | *B* | *C* | *D* |
|  |  |  |  |  |  |  |  |
| *weights* | -2.2 | -1.81 | -.05 | .00 | -.15 | .00 | .16 |
| *error rates* | .12 | .20 | .14 | .37 | .23 | .25 | .36 |

表I.该表显示了用于组合顶部行中的最佳匹配性能的算法相似性估计的权重向量。 大的绝对值表示最有用的预测指标，在这种情况下，是NJIT算法[10]和Viisage [12]。 最后一行给出了单独算法的分类错误比例。 融合将最佳算法的错误率从0.12降低到0.559。

用1至5个因子进行分类的误差率分别为.067，.075，.059，.067和.083。 这些误码率都低于任何一种算法单独运行的最小误码率（参见表1中每种算法的误码率）。 具体来说，数据表明融合，在用PLS回归得出的最佳加权之后，将最佳执行算法（NJIT [10]，误差率为.12）的误差率降低2倍。

为了解释的目的，用于组合相似性得分的权重出现在表I中。这些权重用于组合来自七种算法的相似性得分，以实现匹配和不匹配的脸对分布之间的最大分离。 具有较大绝对值的权重算法对于提高融合性能最有用。

TABLE II

WEIGHT MATRIX FOR HUMAN-ALGORITHM FUSION



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *Human* | *NJIT* | *Viisage* | *CMU* | *B* | *D* |
|  |  |  |  |  |  |  |
| *weights* | .47 | -1.29 | -.71 | -.03 | -.12 | .20 |

表二。 该表显示了用于组合人类和算法相似度估计以实现最佳匹配性能的加权向量。 算法A和C的权重为零，不包含在表中。 将人类添加到PLS中，将算法融合的错误率从0.059降低到人类算法融合的0.008。

以此作为解释指南，很明显，大多数精度的提高来自于两种算法，即NJIT [10]和Viisage [12]，其权重具有最大的绝对值。 这可能是由于这些算法具有计算面部相似性的最大多样化策略。 这种解释看起来可能是因为CMU算法[11]比Viisage算法[12]表现得更好。 因此，组合性能较差的算法可以获得更多的好处，这些算法以不同的方式运行，而不是组合更高性能的类似算法。

B.实验2 - 通过PLS进行人和算法融合

融合人类和算法是否可以增加匹配估计的准确性，并进一步改进分类与融合算法所获得的分类相比较？ 在这个实验中，我们在PLS模型中增加了人类相似性估计。 分析如前所述进行，但使用包含附加到预测矩阵的平均人类相似性数据的列向量。

再次，我们改变了我们保留的PLS因子的数量从1到5.在这种情况下，我们发现双因子解决方案是最强大的，使用前面描述的折叠刀程序。 表II中显示了用于组合人类和算法相似性估计的权重。 通过五个因素的一个因素的性能分别产生0.042,0.008,0.333和0.042的分类错误率。

这些结果表明，当将人类添加到预测矩阵中时，可以获得接近完美的分类。 这表明人类用于分配面部相似性的策略有助于最佳算法所使用的人员。

从以前的工作[9]中值得注意的是，发现人类的准确度低于NJIT [10]，CMU [11]和Viisage [12]，但高于算法A，B，C的准确度和D.在该研究中，来自单个受试者的相似性评分在120个面部对中被折叠以创建每个受试者的ROC曲线。然后对这些单独的ROC曲线进行平均以给出人类准确度的总体估计。在这里，我们对120个脸部对的相似性评分进行了平均，并在各个主题中折叠。有趣的是，尽管可能并不奇怪，但我们发现通过对49个人类个体对每个人脸对的面部相似度进行平均估计，人类错误率为0.12，与NJIT（最佳算法）相当。这表明像算法这样的个体主体可以采用不同的策略来判断面部对的相似性。因此，通过融合对个体受试者的相似性估计进行结合同样可以获得准确性。

# 讨论

人脸识别和算法融合是人脸识别研究人员和公司希望将其系统应用于实际应用的合理目标。 了解算法和人类自身的准确程度是尝试估算算法和人类组合的工作方式的开始。 但是，针对个别算法和人类的准确度的定量测量不足以指导混合系统的开发。 本研究表明算法和人类最有用的融合可能来自将人脸识别系统（算法或人类）与多种人脸识别策略相结合。

在本文中，我们演示了融合算法和人类在难度较大的面部匹配任务中大幅提高了性能。 使用PLS回归来融合算法和人类也产生了如何将融合的各个组成部分最优化组合的精确指示。 这个权重矢量同时作为融合系统的一个配方，并作为算法和人类策略相似性的一个指标来进行相似性判断。

鉴于算法和人都不能在不受控制的环境中完美地执行人脸识别，并且大多数应用程序在循环中都有人工操作员，研究人员合理的目标应该是设计最佳地将算法和人类组合起来的人脸识别策略。 因此，融合算法和人类来创建优秀的混合体能够成为改善重要应用中的人脸匹配性能的有用且实用的方法。

参考文献

1. R. Gross, S. Baker, I. Matthews, and T. Kanade, ˝Face recognition across pose and illumination, ˝ in Handbook of Face Recognition, S. Z. Li and A.K. Jain, Eds. Springer, pp. 193-216, 2005
2. P.J. Phillips, H. Moon, P. Rizvi, and P. Rauss, ˝The FERET evaluation method for face recognition algorithms, ˝ IEEE Transactions on Pattern Analysis and Machine Learning, vol. 22, pp. 1090-1104, 2000.
3. P.J. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and J.M. Bone, ˝FRVT 2002 evaluation report, ˝ Tech. Rep. NISTIR 6965 http://www.frvt.org, 2003.
4. W.J. Braje, ˝Illumination encoding in face recognition: effect of position shift, ˝ Journal of Vision, vol. 3, pp. 161-170, 2003.
5. W.J. Braje, D. Kersten, M.J. Tarr, and N.F. Troje, ˝Illumination effects in face recognition, ˝ Psychobiology, vol. 26, pp. 371-380, 1999.
6. W.J. Braje, G.E. Legge, and D. Kersten, ˝Invariant recognition of natural objects in the presence of shadows, ˝ Perception, vol. 29, pp. 383-398, 2000.
7. P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, ˝Overview of the face recognition grand challenge, ˝ Proc. IEEE Computer Vision & Pattern Recognition, vol. 1, pp. 947-954, 2005.
8. P.J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, W. Worek, ˝Preliminary Face Recognition Grand Challenge Results,˝ in Proceedings of the Seventh International Conference on Automatic Face and Gesture Recognition, 15-24, 2006.
9. A. O’Toole, P.J. Phillips, F. Jiang, J. Ayyad, N. Pénard, and H. Abdi, ˝Face recognition algorithms surpass humans,˝ Techh. Rep. NISTIR, http://face.nist.gov, 2005.
10. C. Liu, ˝Capitalize on dimensionality increasing techniques from improving face recognition Grand Challenge performance, ˝ IEEE Transactions on Pattern Analysis and Machine Learning, 2006, in press.
11. C.M. Xie, M. Savvides, and V. Kumar, ˝Kernel correlation filter based redundant class-dependence feature analysis (KCFA) on FRGC2.0 Data, ˝ IEEE International Workshop Analysis & Modeling Faces & Gestures, pp. 32-43, 2005.
12. M, Husken, B. Brauckmann, S. Gehlen, and C. von der Malsburg, ˝Strategies and benefits of fusion of 2D and 3D face recognition, ˝

Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 3, pp. 174, 2005

1. P. Grother, "Face Recognition Vendor Test 2002 Supplemental Report," Tech. Rep. NISTIR 7083 http://www.frvt.org, 2004.
2. O. Melnik, Y. Vardi, C.-H. Zhang, “Mixed Group Ranks: Preference and Confidence in Classifier Combination,” IEEE Transactions on Pattern Analysis and Machine Learning, vol. 26, pp. 973-981, 2004.
3. J. Czyz, J. Kittler, L. Vanderdorpe, “Combining face verification experts”, In Proceedings 16th International Conference on Pattern Recognition II pp. 28-31, 2002.
4. H. Abdi Partial least squares regression (PLS-regression). In M. Lewis Beck, A. Bryman, T. Futing (Eds.) Encyclopedia for Research Methods for the Social Sciences. Thousand Oaks: CA, Sage, 2003. pp. 792-795.
5. T. Naes, T. Isaksson, T. Fearn and Davis T. Multivariate calibration and classification. Chisester (UK) NIR Publications. 2004.
6. H. Martens, and M. Martens. Multivariate analysis of quality. London: J. Wiley. 2001.
7. A.R. McIntosh, and N. Lobaugh. Partial least squares analysis of neuroimaging data: applications and advances. Neuroimage, vol. 23, pp. 250-263. 2004.
8. H. Abdi Partial least squares regression. In N. J. Salkind (Ed.)

Encyclopedia of measurement and statistics. Thousand Oaks: CA, Sage, 2007. pp. 740-744.

1. H. Abdi Multivariate analysis. In M. Lewis Beck, A. Bryman, T. Futing (Eds.) Encyclopedia for Research Methods for the Social Sciences. Thousand Oaks: CA, Sage, 2003. pp. 699-702.

致谢

这项工作得到了TSWG公司的A. O'Toole和H. Abdi的合同的支持。 P. J. Phillips的部分资助来自国家司法研究所。这项工作是根据“边界安全法”第303条为司法部进行的，编纂为8 U.S.C. 1732.本报告中确定的特定硬件和软件产品用于执行本文档中描述的评估。这种标识在任何情况下都不意味着国家标准与技术研究院的推荐或认可，也不意味着所确定的产品和设备必定是最适合此目的的。 FRGC的主要目标是鼓励和促进人脸识别算法的开发。为了向人脸识别研究团体提供最新算法的公正评估，研究小组自愿将原型实验的相似性分数提交给美国国家标准与技术研究院（NIST）进行分析。NIST分析的结果是匿名的，除非参与算法开发者另有约定。 所有参与团队都可以选择保留匿名或在本报告中被识别。 2005年1月以来，所有算法的性能结果均为2005年1月以来的结果，2005年8月结果为Xie等人。

1. 附录：PLS回归

在本附录中，我们简要介绍了偏最小二乘回归（PLSR），更完整的介绍可以在[16,20]中找到。 MATLAB程序可以从www.utdallas.edu/ herve下载。 PLSR概括并结合主成分分析和多元回归的特征。 其目标是从一组预测变量中最优地预测一组因变量。 具体而言，PLSR搜索一组用于同时分解X和Y的组件（称为潜在向量），其约束条件是这些组件尽可能地解释X和Y之间的协方差。此步骤之后是一个回归步骤 X的分解用于预测Y.

1）符号

由K个因变量描述的I观测值被存储在表示为Y的I K矩阵中，预测器的I矩阵表示为X.不失一般性，X和Y都被假设为居中和归一化。 常见的一组（正交）

将潜在向量存储在I L矩阵T中（即TTT = 1）。 PLRS将X分解为

X = TPT。

其中P是一个叫做X加载矩阵的矩阵。 矩阵Y估计为

Y = TBCT

其中B是以“回归权重”作为对角元素的对角矩阵，C是因变量的“权重矩阵”

2）潜在矢量，负载和权重的计算

通过找到两组权重w和c来获得潜在向量，以便创建（分别）X和Y的列的线性组合，使得它们的协方差是最大的。 具体而言，目标是获得第一对向量

t = Xw和u = Yc。

在约束下

wTw = 1，tTt = 1，tTu最大。

当找到第一个潜在向量时，它将从X和Y中减去，并重复该过程直到X变为空矩阵（有关更多信息，请参阅算法部分）。

3）算法

PLSR的不同组成部分可以通过一系列的奇异值分解和紧缩来找到。 具体来说，第一权向量w和c分别是

矩阵XTY的第一左右奇异向量。 然后使用方程式1导出矢量t和u。用这些矢量表示数值b的计算公式为b = tTu，然后用来从t预测Y为Y = btcT。 X的因子载荷计算为p = Xt。 现在如下从X和Y中减去（即部分输出）t的效应

X = X-tpT，Y = Y-btcT。 然后将向量t，u，w，c和p存储在对应的矩阵中，并将标量b存储为B的对角元素。如果X是空矩阵，则已经找到整个潜在向量集合 ，否则该过程被重新迭代。

4）预测因变量

使用多元回归公式来预测因变量

Y = TBCT = XBPPLS

同

BPLS= PT+ BCT

（其中PT +是PT的Moore-Penrose伪逆）。