基于多媒体应用的无监督学习与监督学习方法比较研究

1st AmiteshKumarSingam*IEEE*计算智能学会 电气与电子工程师协会海得拉巴，印度 0000-0002-2532-0989

2ndBenny L¨ovstr¨om布莱金厄理工学院卡尔斯克鲁纳，瑞典 0000-0003-3824-0942

3rdWlodek J. Kulesza布莱金厄理工学院卡尔斯克鲁纳，瑞典 0000-0001-8351-8830

arXiv:2303.02446v1 [cs.CV] 4 Mar 2023

摘要—在移动通信领域，一些视频应用促进了人们对稳健视频质量评估方法的兴趣。在所有现有方法中，我们更倾向于无参考视频质量评估，因为在参考视频仅部分可用的情况下，这种方法最为必要。我们的研究重点在于基于人类视觉特征，将有效特征构建并融合到一个模型中。本文对监督学习与无监督学习方法进行了比较研究。因此，我们采用支持向量回归算法，将其作为基于无参考的视频质量指标（VQM），利用简化的输入特征进行质量估计。我们得出结论：即使经过降维，我们提出的模型对于 SSIM 质量指标的客观评分仍表现出稀疏性。

索引词—视频质量指标、无参考视频质量指标、支持向量机、主成分分析、结构相似性。

I. 引言

移动设备上的消费级视频应用使用量大幅增加。服务提供商和应用开发者之间竞争激烈，都希望提供更优质的服务；与此同时，人们对先进视频质量评估方法的需求也日益迫切。移动设备或手持设备上的多媒体服务，例如视频聊天和实时视频流，由于参考视频并非总是可用，进一步激发了人们对无参考客观视频质量评估的兴趣。视频服务的一项主要特征是终端用户所感知的服务质量。例如，在传输过程中，视频的视觉质量可能会在采集、加载、保存或播放时下降，而视频缓冲过程中也可能在上述任一阶段出现失真。由于视频服务的终端用户是视频质量的最终评判者，因此，从技术角度来说，观众验证视频质量的过程称为主观视频质量评估。然而，主观实验往往耗时过长、操作不便且成本高昂，而且必须遵循特定建议才能获得标准化结果。基于这些原因，人们需要一些智能方法，在保证准确性的同时，快速、经济地预测视频质量评估结果。

II. 视频质量评估

在我们的研究中，我们严格采用了 ITU-R BT 500-10 中推荐的具体规范（如 [1] 所述），以及 [2] 中的 VQEG混合测试计划，相关内容将在后续章节中简要说明。

III. 提出的方法

所提出的方法主要包括提取与视觉质量相关的码流参数，并构建基于机器学习的质量预测模型。这些参数经过精心挑选，以控制复杂度，并获取能够表征编码失真的合理编码信息。以下将介绍提取的参数，以及将这些参数纳入所提出模型的理由。

IV. H.264 比特流数据的特征提取

已使用JM参考软件16.0版对编码视频序列进行了特征提取。对编码过程后生成的跟踪文件中的比特流数据进行了特征提取。

1）比特率。2）帧率。3）大小为16x16的帧间宏块百分比。4）大小为4x4的帧间宏块百分比。5）大小为8x8的帧间宏块百分比。6）平均量化参数。

V. 无监督学习方法

通常，在处理高维数据时，增加更多特征会影响系统或模型的性能，并提高系统的复杂度。因此，为了克服维度灾难，需要考虑效率、分类性能和建模便捷性等因素。效率又可分为测量成本、存储成本和计算成本。尤其是在机器学习中，在高维诱导空间中，观测值数量有限，而每个输入数据都可能具有多种取值，因此任何模型都需要大量特征进行训练。

这样才能确保有足够多的样本，涵盖各种取值排列。因此，我们的核心思路是在尽可能少丢失信息的前提下降低输入数据的维度。降维适用于数据可视化、获得紧凑表示以及减少计算处理。此外，减少维度数量还可以将包含重要信息的特征与重要性较低的特征区分开来，从而进一步洞察数据的本质，否则这些信息可能难以发现。降维技术有多种，例如偏最小二乘（PLS）、主成分分析（PCA）、奇异值分解（SVD）、核主成分分析（KPCA）、因子分析和层次聚类等。PLS 和 PCA 在回归分析中有两项作用：通过线性变换将高度相关的变量转换为相互独立的变量，以及进行降维。在回归问题中，因此我们在本研究中选择使用 PCA 进行降维。

A. 降维

a)：尤其是在机器学习领域，对于高维诱导空间中的有限数量观测值，且每个输入数据都有大量可能取值的情况，任何模型都需要大量特征进行训练，以确保有足够多的样本涵盖各种值的排列。因此，我们的核心思路是在尽可能少丢失信息的前提下减少输入的维度。降维适用于数据可视化、获得紧凑表示以及降低计算负载。此外，减少维度数量可以将包含重要数据的特征与较不重要的特征区分开来，从而进一步洞察数据的本质，而这些本质在其他情况下可能无法被发现。降维技术有多种，例如主成分分析、奇异值分解、核主成分分析、因子分析和层次聚类等。在我们的研究中，我们采用 PCA 进行降维。

B. 主成分分析

主成分分析是一种统计方法，采用正交线性变换技术，将一组向量投影到新的维度空间中；该空间的分量彼此线性不相关，并按方差从大到小排列。通常认为，投影空间中最前面的坐标包含最重要的数据，且具有较大的方差。减少输入向量的维度有助于提升泛化性能。

a) : 主成分分析用于降低我们提出的模型的输入空间维度。降低输入空间的维度能够降低系统的复杂度，并缩短模型训练所需的时间。主成分分析可用作数据预处理方法，从数据中提取不相关特征。

b): 计算主成分分析：如果 X 是一个具有 m 个维度的输入向量，为了降低维度，我们需要从给定的输入向量中剔除不重要的数据。

$$\vec{X} = [x_1, x_2, .., x_{m0}, x_{m0+1}, ..., x_m] \tag{1}$$

c): 我们必须将输入向量 X 映射到其他维度X_{new}。通过设计一个具有 m 行和 m 列的变换矩阵 (T)，并将矩阵 T 与输入向量相乘，我们得到X_{new}。因此，表达式如下：

$$\vec{X}_{new} = T.\vec{X} \tag{2}$$

- X 是一个均值为零的 m 维随机向量。
- $\vec{X}_{new}$ 是新的空间维度。

d) : 主成分分析是一种多变量技术，用于降低具有大量相关变量的输入特征的维度。因此，为了降低输入向量的维度，我们需要最大限度地提高方差的下降率。主成分分析的数学模型如下：

e) : 让输入随机向量，q 是一个 m 维单位向量。根据向量 $\vec{q}$的性质，$\vec{q}$ 的欧几里得范数为 1。$\| \vec{q} \| = 1$ 将 X 投影到投影空间 A 中的 q 上，可得：

$$A = \vec{X}^T \vec{q} \tag{3}$$

方差

$$E[A^2] = \vec{q} R_{xx} \vec{q}^T \tag{4}$$

其中 R_{xx} 是相关矩阵

$$R_{xx} = E[\vec{X}\vec{X}^T] \tag{5}$$

令 $\psi(\vec{q})$为方差探针

$$\psi(\vec{q}) = \vec{q} R_{xx} \vec{q}^T \tag{6}$$

对于方差的最小值

$$\psi(\vec{q} + \delta\vec{q}) = \psi(\vec{q}) \tag{7}$$

计算协方差矩阵 R 的特征值 λ_1，λ_2，$.....\lambda_m$ 以及对应的正交特征向量 $\vec{q}_1$，$\vec{q}_2$，$.....\vec{q}_m$，并按其大小排列 $\lambda_1 \geq \lambda_2 \geq ..... \geq \lambda_m$

$$R\vec{q}_j = \lambda_j \vec{q}_j \tag{8}$$

其中λ_1 是所有特征值中的最大值。

$$\vec{Q} = [\vec{q}_1, \vec{q}_2, ...\vec{q}_j, ..., \vec{q}_m] \tag{9}$$

$$\Delta = diagonal[\lambda_1, \lambda_2, ...\lambda_m] \tag{10}$$

Q是满足条件的正交矩阵

$$\vec{q}_i{}^T \vec{q}_j = \begin{matrix} 1, \text{if } i = j \\ 0, \text{if } i \neq j \end{matrix} \tag{11}$$

正交相似变换表示为

$$\vec{Q}^T R \vec{Q} = \Delta \tag{12}$$

展开形式为

$$\vec{q_j}^T R\vec{q_k} = \begin{array}{l}\lambda_j, \text{if } k=j \\ 0, \text{if } k\neq j\end{array} \tag{13}$$

相关矩阵 R 可用其特征值和特征向量表示如下

$$R = \sum_{j=1}^{m} \lambda_j \vec{q_j} {[\vec{q}]_j}^T \tag{14}$$

$$\psi(\vec{q_j}) = \lambda_j \tag{15}$$

$$a_j = X^T \vec{q_j} \tag{16}$$

对于 j=1,2,....m（表示分析），a_j=将 X 投影到主方向上，

$$X = \sum_{j=1}^{m} a_j \vec{q_j} \tag{17}$$

对于 j=1,...., m（表示综合），设特征值$\lambda_1, \lambda_2, ...\lambda_l$ 是相关矩阵 R_{xx} 中最大的值，并且

$$\hat{X} = \sum_{j=1}^{l} a_j \vec{q_j} \tag{18}$$

$\hat{X}$ 并非精确值，而是截断至 l 项后得到的近似解，但向量 X 的维度得以保留。

- PCA 编码器和解码器的矩阵表示

  are

  f) :
  $$\begin{bmatrix} x_1 \\ | \\ x_m \end{bmatrix} \rightarrow \begin{bmatrix} [\vec{q_1}]^T \\ | \\ [\vec{q_l}]^T \end{bmatrix} \rightarrow \begin{bmatrix} a_1 \\ | \\ a_l \end{bmatrix}$$

  g) :
  $$\begin{bmatrix} a_1 \\ | \\ a_l \end{bmatrix} \rightarrow \begin{bmatrix} \vec{q_1} & --\vec{q_l} \end{bmatrix} \rightarrow \begin{bmatrix} \hat{x_1} \\ | \\ \hat{x_m} \end{bmatrix}$$

$$\vec{e} = X - \hat{X} = \sum_{i=l+1}^{m} a_j \vec{q_j} \tag{19}$$

$\hat{}$ 误差与 X 正交，因为 PCA 将线性输入空间转换为正交空间 $\vec{e}\hat{}x = 0$

- 输入数据的 m 个分量的总方差 $\vec{X}$

$$\sum_{j=1}^{m} {\sigma_j}^2 = \sum_{j=1}^{m} \lambda_j \tag{20}$$

- X的 l 个分量的总方差 $\hat{}$

$$\sum_{j=1}^{l} {\sigma_j}^2 = \sum_{j=1}^{l} \lambda_j \tag{21}$$

- 误差向量中输入数据向量的 (l-m) 个分量的总方差

$$\sum_{j=l+1}^{m} {\sigma_j}^2 = \sum_{j=l+1}^{m} \lambda_j \tag{22}$$

其中λ_{l+1}， λ_{l+2}， ....λ_m 对应较小的方差。因此，主成分分析检测到的输入向量的方差取决于特征值。

本节推导了从 2.8 到 2.29 的所有方程

h) : λ_i 按降序排列，并由主成分 (l) 解释的方差比例为

$$\frac{\lambda_1 + \lambda_2 + ... + \lambda_l}{\lambda_1 + \lambda_2 + ... + \lambda_l + ... + \lambda_m}$$

如果各维度高度相关，那么特征值较大的特征向量数量就会较少，因此其维度为 l≪m，并且可以实现大幅降维。在图像和语音处理任务中，我们发现了这类典型情况，因为输入高度相关。如果各维度互不相关，那么主成分分析就不会带来任何收益。

VI. 监督学习方法

基于核的学习方法分为监督学习和无监督学习算法。核方法通过线性或非线性映射，将输入数据集映射到高维特征空间中，这一过程也被称为核技巧，从而解决各种问题。近年来，研究人员提出了几种强大的基于核的模型，例如支持向量机、核费舍尔判别和核主成分分析，这些模型可用于回归、分类、降维及其他任务。在我们的研究工作中，我们采用支持向量机（SVM）算法进行回归分析。

A. 支持向量机

支持向量机（SVM）是一种由 Vladimir VaDnik [3] 发明的监督式强大学习算法，通常用于分类和回归分析。其公式建立在结构风险最小化原则之上，该原则包含容量控制，以防止传统神经网络等基于经验风险最小化原则的学习算法出现过拟合问题。在我们的研究工作中，我们进行了回归分析，其中支持向量回归通过映射从 H.264 码流数据中提取的简化输入特征，并将其与期望值或真实值进行回归，以非常有效的方式展现了机器学习的优势，以及学习复杂数据模式的能力。SVM 的工作机制是通过非线性映射，将非线性输入数据映射到高维核诱导空间中，从而在核空间内求解线性方程组 [**?**]。SVM 中引入了一种不敏感损失函数，用于衡量风险；核函数具有足够的灵活性，使 SVM 能够搜索多种假设空间 [4]。

B. 回归分析

支持向量回归模型能够在无参考视频质量评价中更准确地预测视频质量，并与人类视觉系统相一致。支持向量回归是一种监督学习方法，它通过非线性映射将输入数据映射到高维核诱导特征空间中。因此，非线性函数被转换为线性函数，并在高维空间中进行回归。模型的容量由不依赖于特征空间维度的参数控制。在支持向量回归中，

引入了一种称为 epsilon 的损失函数，如图所示，它忽略真实值所在区域内的误差。

a)：给定的输入数据集 X 首先通过非线性映射映射到一个 m 维特征空间，然后在该 m 维特征空间中构建一个线性模型。特征空间中的线性模型表示为

$$f(x,\omega)=\sum_{j=1}^{m}\omega_j g_j(x)+b \tag{23}$$

$g_j(x)$ $j=1$，$...m$ 是非线性变换的函数，而项 b 是偏置。由于通过对输入数据进行预处理，假定其均值为零，因此省略偏置。带有松弛变量的 ε带如下图所示。

松弛变量衡量训练点上误差的代价，而对于 ε带内的所有点，ε误差为零。支持向量回归引入了一种称为 ε-不敏感的损失函数。质量估计可以借助该损失函数进行衡量。

$$L_\varepsilon(y,f(x,\omega))=\begin{cases}0, & \text{if } |y-f(x,\omega)|\le\varepsilon,\\ |y-f(x,\omega)|-\varepsilon, & \text{otherwise.}\end{cases} \tag{24}$$

经验风险定义为

$$R_{emp}(\omega)=\frac{1}{n}\sum_{i=1}^{n}L_\varepsilon(y_i,f(x_i,\omega)) \tag{25}$$

线性回归通过支持向量回归在高维核诱导特征空间中采用 ε- 不敏感损失来执行，同时通过最小化来降低模型的复杂度。在 ε-不敏感区域之外，通过引入松弛变量（ξ_i，ξ_i^*，其中 i=1，....，n）来衡量训练数据样本的偏差。松弛变量为非负值，用于确定上界和下界。因此，支持向量回归被表述为一个最小化问题。

$$min\frac{1}{2}\,||\omega||^{\,2}+C\sum_{i=1}^{n}(\xi_i+\xi_i^{*}) \tag{26}$$

$$s.t\begin{cases}y_i-f(x_i,\omega)\le\varepsilon+\xi_i^*\\ f(x_i,\omega)\text{–}y_i\le\varepsilon+\xi_i\\ \xi_i,\xi_i^*\ge 0, \qquad \text{i=1,....,n.}\end{cases}$$

该优化问题可以转化为对偶问题

$$f(x)=\sum_{i=1}^{n_{sv}}(\alpha-\alpha_i^*)K(x_i,x)) \tag{27}$$

$$w.r.t0\le\alpha_i^*\le C,0\le\alpha_i\le C$$

b)：其中，n_{sv}- 支持向量数量，α,α_i^* 是某些样本的拉格朗日系数。其中一些系数的均值非零，对应的样本即为支持向量，核函数为

$$K(x,x_i)=\sum_{j=1}^{m}g_j(x)g_j(x_i) \tag{28}$$

常用的核函数有：线性核函数、多项式核函数、径向基函数核。在我们提出的模型中，我们使用径向基函数核作为核函数

$$K(y,y_i)=exp\frac{-||y-y_i||^2}{\sigma^2} \tag{29}$$

本节推导方程 2.41-2.47。其中，σ 为核参数。支持向量回归的泛化性能取决于对核及元参数（C、 ε）的合理设置。由于支持向量回归模型的复杂度取决于全部三个参数，因此最优参数选择问题也变得复杂。支持向量机将元参数视为用户为回归指定的输入。核类型和核函数参数的选择取决于应用领域知识，同时也应反映输入（x）训练数据的分布。C 决定模型复杂度与可容忍偏差超过 ε 的程度之间的权衡。若 C 过大或趋于无穷，则目标是在优化公式中不考虑模型复杂度部分的情况下，最小化经验风险。ε 控制 ε 不敏感区间的宽度，以拟合训练数据。该参数的取值会影响用于构建回归函数的支持向量数量：若 ε 较大，则只会选取少量支持向量；较大的 ε 值会产生更平坦的估计值。因此，C 和 ε 值都会影响模型复杂度，但方式不同。

VII. 测试方法

我们采用 k 折交叉验证（交叉验证）策略来训练和验证所提出的模型。交叉验证（CV）是一种用于评估预测模型性能的技术。通常，交叉验证是一种重采样策略，通过对可用数据进行随机子采样来验证所提出模型的性能。换言之，原始数据被随机划分为 k 折。在每一轮交叉验证中，使用（k – 1）折进行训练，其余一折用于验证；这一过程重复 k 次，确保每一折恰好用于一次模型验证。所提出模型的性能估计值由 k 轮中各折所得结果的平均值确定 [5]。训练期间会监测验证集上出现的误差。在训练初期，验证误差通常会下降，训练集误差也会如此。

a)：主要来说，视频经网络传输后，在解码器端通过提取比特流数据获得输入特征。我们提出的模型已使用输入特征 (X) 进行训练和测试；这些特征由上述比特流数据中提取的 6 个参数组成，并分别对应 SSIM 质量指标的目标值。

实例数	MAE	MSE	RMSE	R平方
实例 1	0.033981	0.002152	0.046385	0.460293
实例 2	0.030235	0.001535	0.039183	0.614883
实例 3	0.021469	0.000859	0.029302	0.784624
实例 4	0.036298	0.002618	0.051165	0.343333

*b)：*所提模型的统计分析给出了表中所述 4 个实例的平均绝对误差、均方误差、均方根误差和 R 平方（拟合优度度量）值；基于 R 平方来量化所提模型与 SSIM 指标之间关系的强度。

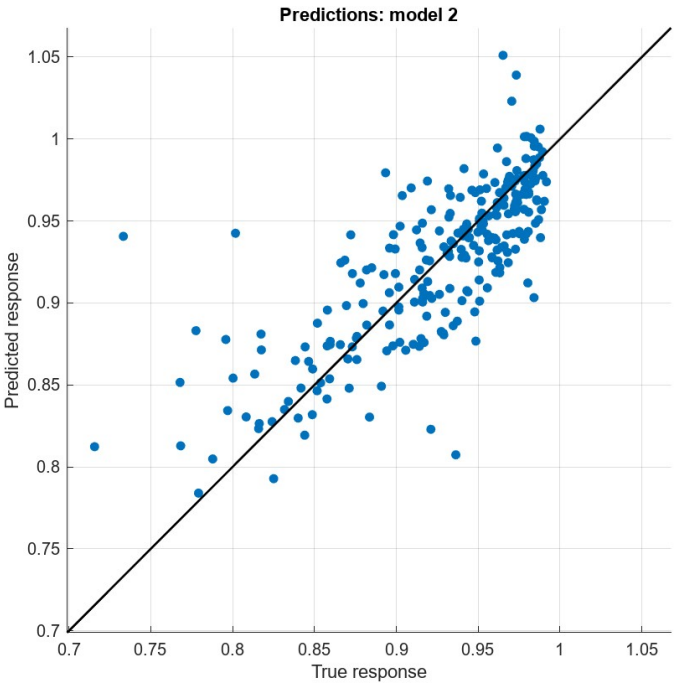


图 1。

上图展示了基于支持向量机，在高维核诱导空间中对输入数据进行线性拟合的结果。

VIII. 结论

我们得出结论：对于 SSIM 质量指标的客观评分，即使经过降维，我们提出的模型仍表现出稀疏性。未来的工作将致力于提高预测准确率，而这取决于是否决定消除所提模型的稀疏性。

参考文献

[1] “ITU-R 无线电通信部门，ITU-R 建议书 BT.500-12”，2009，http://www.itu.int/. [2] “VQEG 混合测试计划，版本 1.2”，2011，ftp://vqeg.its.bldrdoc.gov. [3] Vapnik.V，《统计学习理论》。Wiley，1998。[4] Q. Miao 和 S.-F. Wang，“基于支持向量回归的非线性模型预测控制”，载于《国际机器学习与控制论会议论文集》，第 3 卷，2002 年，第 1657–1661 页，第 3 卷。