



(12) 发明专利申请

(10) 申请公布号 CN 105389326 A

(43) 申请公布日 2016. 03. 09

(21) 申请号 201510591147. 7

G06K 9/62(2006. 01)

(22) 申请日 2015. 09. 16

(71) 申请人 中国科学院计算技术研究所

地址 100190 北京市海淀区中关村科学院南路6号

申请人 北京联索科技有限公司

(72) 发明人 张博 史忠植 王伟 齐保元

马刚

(74) 专利代理机构 北京律诚同业知识产权代理

有限公司 11006

代理人 祁建国 梁挥

(51) Int. Cl.

G06F 17/30(2006. 01)

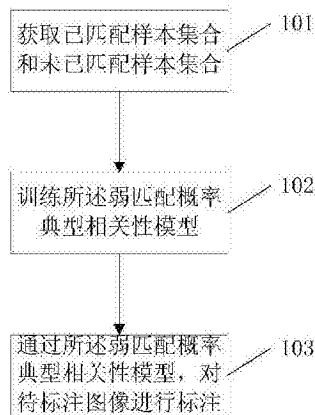
权利要求书3页 说明书11页 附图6页

(54) 发明名称

基于弱匹配概率典型相关性模型的图像标注方法

(57) 摘要

本发明公开了基于弱匹配概率典型相关性模型的图像标注方法及系统,本发明涉及网络跨媒体信息处理技术领域,包括获取图像数据库中已标注图像与未标注图像,分别提取所述已标注图像与所述未标注图像的图像特征和文本特征,生成已匹配样本集合和未匹配样本集合,所述已匹配样本集合包括已标注图像特征集合与已标注文本特征集合,所述未匹配样本集合包括未标注图像特征集合与未标注文本特征集合;根据所述已匹配样本集合与所述未匹配样本集合,训练所述弱匹配概率典型相关性模型;通过所述弱匹配概率典型相关性模型,对待标注图像进行标注。本发明同时使用标注图像及其关键词和未标注图像学习视觉模态和文本模态之间的关联,准确对未知图像进行标注。



1. 一种基于弱匹配概率典型相关性模型的图像标注方法,其特征在于,包括:步骤 1, 获取图像数据库中已标注图像与未标注图像,分别提取所述已标注图像与所述未标注图像的图像特征和文本特征,生成已匹配样本集合和未匹配样本集合,所述已匹配样本集合包括已标注图像特征集合与已标注文本特征集合,所述未匹配样本集合包括未标注图像特征集合与未标注文本特征集合;

步骤 2,根据所述已匹配样本集合与所述未匹配样本集合,训练所述弱匹配概率典型相关性模型;

步骤 3,通过所述弱匹配概率典型相关性模型,对待标注图像进行标注。

2. 如权利要求 1 所述的基于弱匹配概率典型相关性模型的图像标注方法,其特征在于,所述弱匹配概率典型相关性模型的公式为:

$$D = \left\{ \left(\mathbf{x}_1^i, \mathbf{x}_2^i \right) \right\}_{i=1}^{N_p} \cup \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1} \cup \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$$

$$L(\theta) = \prod_{i=1}^{N_p} P\left(\mathbf{x}_1^i, \mathbf{x}_2^i; \theta\right) \prod_{j=N_p+1}^{N_1} P\left(\mathbf{x}_1^j; \theta\right) \prod_{k=N_p+1}^{N_2} P\left(\mathbf{x}_2^k; \theta\right)$$

其中, $D = \left\{ \left(\mathbf{x}_1^i, \mathbf{x}_2^i \right) \right\}_{i=1}^{N_p} \cup \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1} \cup \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$ 表示完整的观察样本集合,包含了匹配和未匹配样本, N_p 为成对观察样本集合的样本数量, $\mathbf{X}_1^{(p)} = \left\{ \left(\mathbf{x}_1^i \right) \right\}_{i=1}^{N_p}$ 和 $\mathbf{X}_2^{(p)} = \left\{ \left(\mathbf{x}_2^i \right) \right\}_{i=1}^{N_p}$, 其中每一个样本 \mathbf{x}_1^i (\mathbf{x}_2^i) 代表一个 m_1 (m_2) 维向量, $\mathbf{X}_1^{(u)} = \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1}$ 与 $\mathbf{X}_2^{(u)} = \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$ 表示未匹配样本集合, 其中 $\mathbf{X}_1^{(u)}$ 与 $\mathbf{X}_2^{(u)}$ 相互独立生成, 假设样本之间相互独立, 其极大似然值 $L(\theta)$, $P\left(\mathbf{x}_1^i, \mathbf{x}_2^i; \theta\right)$ 服从概率典型相关性分析模型, $P\left(\mathbf{x}_1^j; \theta\right)$ 和 $P\left(\mathbf{x}_2^k; \theta\right)$ 分别表示未匹配样本集合 $\mathbf{X}_1^{(u)} = \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1}$ 和 $\mathbf{X}_2^{(u)} = \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$ 的概率分布。

3. 如权利要求 2 所述的基于弱匹配概率典型相关性模型的图像标注方法,其特征在于,所述已匹配样本集合通过概率典型相关性分析模型计算 \mathbf{x}_1^i 和 \mathbf{x}_2^i 。

4. 如权利要求 2 所述的基于弱匹配概率典型相关性模型的图像标注方法,其特征在于,对应所述未匹配样本集合 $\mathbf{X}_1^{(u)} = \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1}$ 和 $\mathbf{X}_2^{(u)} = \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$, \mathbf{x}_1^j 和 \mathbf{x}_2^k 通过以下公式获得:

$$P\left(\mathbf{x}_1^j; \theta\right) = \int P\left(\mathbf{x}_1^j | \mathbf{z}_1^j\right) P\left(\mathbf{z}_1^j\right) d\mathbf{z}_1^j \sim N\left(\mu_1, \mathbf{W}_1 \mathbf{W}_1^T + \psi_1\right)$$

$$P\left(\mathbf{x}_2^k; \theta\right) = \int P\left(\mathbf{x}_2^k | \mathbf{z}_2^k\right) P\left(\mathbf{z}_2^k\right) d\mathbf{z}_2^k \sim N\left(\mu_2, \mathbf{W}_2 \mathbf{W}_2^T + \psi_2\right)$$

其中对于未匹配样本集合 $\mathbf{X}_1^{(u)} = \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1}$ 和 $\mathbf{X}_2^{(u)} = \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$, \mathbf{x}_1^j 和 \mathbf{x}_2^k 则分别由隐变量 \mathbf{z}_1^j 和 \mathbf{z}_2^k 通过线性变换 \mathbf{W}_1 和 \mathbf{W}_2 附加高斯噪声 ε_1 和 ε_2 获得, ψ_1 , ψ_2 , μ_1 和 μ_2 分别表示随机变量 \mathbf{x}_1 和 \mathbf{x}_2 观察样本集合的协方差和均值。

5. 如权利要求 2 所述的基于弱匹配概率典型相关性模型的图像标注方法,其特征在于,通过以下公式获取所述弱匹配概率典型相关性模型中成对样本的投影:

$$E\left(\mathbf{z}_1^j | \mathbf{x}_1^j\right) = \mathbf{W}_1^T \left(\mathbf{W}_1 \mathbf{W}_1^T + \psi_1 \right)^{-1} \left(\mathbf{x}_1^j - \mu_1 \right)$$

$$E(z^t | x_2^t) = \hat{W}_2^T (\hat{W}_1 \hat{W}_2^T + \hat{\psi}_2)^{-1} (x_2^t - \hat{\mu}_2)$$

其中, $E(z^t | x_1^t)$, $E(z^t | x_2^t)$ 构成了成对样本 (x_1^t, x_2^t) 在 SemiPCCA 隐空间的典型投影, \hat{W}_1 和 \hat{W}_2 表示线性变换矩阵, ε_1 和 ε_2 表示高斯噪声, $\hat{\psi}_1$, $\hat{\psi}_2$, $\hat{\mu}_1$ 和 $\hat{\mu}_2$ 分别表示随机变量 x_1 和 x_2 观察样本集合的协方差和均值。

6. 一种基于弱匹配概率典型相关性模型的图像标注系统, 其特征在于, 包括: 获取已匹配样本集与未匹配样本集模块, 用于获取图像数据库中已标注图像与未标注图像, 分别提取所述已标注图像与所述未标注图像的图像特征和文本特征, 生成已匹配样本集合和未匹配样本集合, 所述已匹配样本集合包括已标注图像特征集合与已标注文本特征集合, 所述未匹配样本集合包括未标注图像特征集合与未标注文本特征集合;

训练模型模块, 用于根据所述已匹配样本集合与所述未匹配样本集合, 训练所述弱匹配概率典型相关性模型;

标注模块, 用于通过所述弱匹配概率典型相关性模型, 对待标注图像进行标注。

7. 如权利要求 6 所述的基于弱匹配概率典型相关性模型的图像标注方法, 其特征在于, 所述弱匹配概率典型相关性模型的公式为:

$$D = \left\{ (x_1^t, x_2^t) \right\}_{t=1}^{N_p} \cup \left\{ (x_1^t) \right\}_{t=N_p+1}^{N_1} \cup \left\{ (x_2^t) \right\}_{t=N_p+1}^{N_2}$$

$$L(\theta) = \prod_{t=1}^{N_p} P(x_1^t, x_2^t; \theta) \prod_{t=N_p+1}^{N_1} P(x_1^t; \theta) \prod_{t=N_p+1}^{N_2} P(x_2^t; \theta)$$

其中, $D = \left\{ (x_1^t, x_2^t) \right\}_{t=1}^{N_p} \cup \left\{ (x_1^t) \right\}_{t=N_p+1}^{N_1} \cup \left\{ (x_2^t) \right\}_{t=N_p+1}^{N_2}$ 表示完整的观察样本集合, 包含了匹配和未匹配样本, N_p 为成对观察样本集合的样本数量, $X_1^{(p)} = \left\{ (x_1^t) \right\}_{t=1}^{N_p}$ 和 $X_2^{(p)} = \left\{ (x_2^t) \right\}_{t=1}^{N_p}$, 其中每一个样本 x_1^t (x_2^t) 代表一个 m_1 (m_2) 维向量, $X_1^{(u)} = \left\{ (x_1^t) \right\}_{t=N_p+1}^{N_1}$ 与 $X_2^{(u)} = \left\{ (x_2^t) \right\}_{t=N_p+1}^{N_2}$ 表示未匹配样本集合, 其中 $X_1^{(u)}$ 与 $X_2^{(u)}$ 相互独立生成, 假设样本之间相互独立, 其极大似然值 $L(\theta)$, $P(x_1^t, x_2^t; \theta)$ 服从概率典型相关性分析模型, $P(x_1^t; \theta)$ 和 $P(x_2^t; \theta)$ 分别表示未匹配样本集合 $X_1^{(u)} = \left\{ (x_1^t) \right\}_{t=N_p+1}^{N_1}$ 和 $X_2^{(u)} = \left\{ (x_2^t) \right\}_{t=N_p+1}^{N_2}$ 的概率分布。

8. 如权利要求 7 所述的基于弱匹配概率典型相关性模型的图像标注方法, 其特征在于, 所述已匹配样本集合通过概率典型相关性分析模型计算 x_1^t 和 x_2^t 。

9. 如权利要求 7 所述的基于弱匹配概率典型相关性模型的图像标注方法, 其特征在于, 对应所述未匹配样本集合 $X_1^{(u)} = \left\{ (x_1^t) \right\}_{t=N_p+1}^{N_1}$ 和 $X_2^{(u)} = \left\{ (x_2^t) \right\}_{t=N_p+1}^{N_2}$, x_1^t 和 x_2^t 通过以下公式获得:

$$P(x_1^t; \theta) = \int P(x_1^t | z_1^t) P(z_1^t) dz_1^t \sim N(\mu_1, W_1 W_1^T + \psi_1)$$

$$P(x_2^t; \theta) = \int P(x_2^t | z_2^t) P(z_2^t) dz_2^t \sim N(\mu_2, W_2 W_2^T + \psi_2)$$

其中对于未匹配样本集合 $X_1^{(u)} = \left\{ (x_1^t) \right\}_{t=N_p+1}^{N_1}$ 和 $X_2^{(u)} = \left\{ (x_2^t) \right\}_{t=N_p+1}^{N_2}$, x_1^t 和 x_2^t 则分别由隐变量 z_1^t 和 z_2^t 通过线性变换 W_1 和 W_2 附加高斯噪声 ε_1 和 ε_2 获得, ψ_1 , ψ_2 , μ_1 和 μ_2 分别表示随机变

量 x_1 和 x_2 观察样本集合的协方差和均值。

10. 如权利要求 7 所述的基于弱匹配概率典型相关性模型的图像标注方法,其特征在于,通过以下公式获取所述弱匹配概率典型相关性模型中成对样本的投影:

$$E(z^i|x_1^i) = \hat{W}_1^T (\hat{W}_1 \hat{W}_1^T + \hat{\psi}_1)^{-1} (x_1^i - \hat{\mu}_1)$$

$$E(z^i|x_2^i) = \hat{W}_2^T (\hat{W}_2 \hat{W}_2^T + \hat{\psi}_2)^{-1} (x_2^i - \hat{\mu}_2)$$

其中, $E(z^i|x_1^i)$, $E(z^i|x_2^i)$ 构成了成对样本 (x_1^i, x_2^i) 在 SemiPCCA 隐空间的典型投影, \hat{W}_1 和 \hat{W}_2 表示线性变换矩阵, ε_1 和 ε_2 表示高斯噪声, $\hat{\psi}_1$, $\hat{\psi}_2$, $\hat{\mu}_1$ 和 $\hat{\mu}_2$ 分别表示随机变量 x_1 和 x_2 观察样本集合的协方差和均值。

基于弱匹配概率典型相关性模型的图像标注方法

技术领域

[0001] 本发明涉及网络跨媒体信息处理技术领域,特别涉及基于弱匹配概率典型相关性模型的图像标注方法。

背景技术

[0002] 物联网、互联网等拥有丰富的文本、图像、视频和音频等多媒体信息资源,这些信息资源是异构的,很难直接发现它们之间的关联,典型相关性分析(Canonical correlation analysis, CCA)是一种用来分析两组随机变量之间相关性的统计分析工具,其相关性保持特征已经在理论上得到证明,应用于经济学、气象和基因组数据分析等领域,CCA通过统计方法找到两组异构多模态特征之间的潜在关系,从底层特征上用统一的模型将不同类型的多模态数据关联起来,同时尽可能地发现和保持数据间潜在的相关性。

[0003] 典型相关性分析中两组相关的随机变量可以来自多种信息来源(如同一个人的声音和图像),也可以是从同一来源的信息中抽取的不同特征(如图像的颜色特征和纹理特征),但训练数据必须一对一严格匹配,很多原因造成这种严格匹配的训练数据难以获得,如多传感器采集系统中传感器采样频率不同步或传感器故障,会造成不同通道采集来的数据不同步或丢失某一通道数据;单模态数据比较容易获得,但人工匹配却非常地费时费力,实际中,面对的多模态数据经常是只有少量一对一严格匹配,其余大量数据未匹配,称之为弱匹配多模态数据。

[0004] 面向弱匹配多模态数据的典型相关性分析有两种基本的方法:丢弃未匹配数据,只使用典型相关性分析处理严格匹配的多模态数据;根据特定准则,匹配多模态数据,但这两种方法都不可能获得理想的结果。

发明内容

[0005] 针对现有技术的不足,本发明提出一种基于弱匹配概率典型相关性模型的图像标注方法及系统。

[0006] 本发明提出一种基于弱匹配概率典型相关性模型的图像标注方法,包括:步骤1,获取图像数据库中已标注图像与未标注图像,分别提取所述已标注图像与所述未标注图像的图像特征和文本特征,生成已匹配样本集合和未匹配样本集合,所述已匹配样本集合包括已标注图像特征集合与已标注文本特征集合,所述未匹配样本集合包括未标注图像特征集合与未标注文本特征集合;

[0007] 步骤2,根据所述已匹配样本集合与所述未匹配样本集合,训练所述弱匹配概率典型相关性模型;

[0008] 步骤3,通过所述弱匹配概率典型相关性模型,对待标注图像进行标注。

[0009] 所述的基于弱匹配概率典型相关性模型的图像标注方法,所述弱匹配概率典型相关性模型的公式为:

$$[0010] \quad \mathbf{D} = \left\{ \left\{ \mathbf{x}'_1, \mathbf{x}'_2 \right\}_{j=1}^{N_p} \cup \left\{ \mathbf{x}'_1 \right\}_{j=N_p+1}^{N_1} \cup \left\{ \mathbf{x}'_2 \right\}_{k=N_p+1}^{N_2} \right\}$$

$$[0011] \quad L(\theta) = \prod_{j=1}^{N_p} P(\mathbf{x}'_1, \mathbf{x}'_2; \theta) \prod_{j=N_p+1}^{N_1} P(\mathbf{x}'_1; \theta) \prod_{k=N_p+1}^{N_2} P(\mathbf{x}'_2; \theta)$$

[0012] 其中, $\mathbf{D} = \left\{ \left\{ \mathbf{x}'_1, \mathbf{x}'_2 \right\}_{j=1}^{N_p} \cup \left\{ \mathbf{x}'_1 \right\}_{j=N_p+1}^{N_1} \cup \left\{ \mathbf{x}'_2 \right\}_{k=N_p+1}^{N_2} \right\}$ 表示完整的观察样本集合, 包含了匹配和未匹配样本, N_p 为成对观察样本集合的样本数量, $\mathbf{X}_1^{(p)} = \left\{ \left\{ \mathbf{x}'_1 \right\}_{j=1}^{N_p} \right\}$ 和 $\mathbf{X}_2^{(p)} = \left\{ \left\{ \mathbf{x}'_2 \right\}_{j=1}^{N_p} \right\}$, 其中每一个样本 \mathbf{x}'_1 (\mathbf{x}'_2) 代表一个 m_1 (m_2) 维向量, $\mathbf{X}_1^{(u)} = \left\{ \left\{ \mathbf{x}'_1 \right\}_{j=N_p+1}^{N_1} \right\}$ 与 $\mathbf{X}_2^{(u)} = \left\{ \left\{ \mathbf{x}'_2 \right\}_{k=N_p+1}^{N_2} \right\}$ 表示未匹配样本集合, 其中 $\mathbf{X}_1^{(u)}$ 与 $\mathbf{X}_2^{(u)}$ 相互独立生成。假设样本之间相互独立, 其极大似然值 $L(\theta)$, $P(\mathbf{x}'_1, \mathbf{x}'_2; \theta)$ 服从概率典型相关性分析模型, $P(\mathbf{x}'_1; \theta)$ 和 $P(\mathbf{x}'_2; \theta)$ 分别表示未匹配样本集合 $\mathbf{X}_1^{(u)} = \left\{ \left\{ \mathbf{x}'_1 \right\}_{j=N_p+1}^{N_1} \right\}$ 和 $\mathbf{X}_2^{(u)} = \left\{ \left\{ \mathbf{x}'_2 \right\}_{k=N_p+1}^{N_2} \right\}$ 的概率分布。

[0013] 所述的基于弱匹配概率典型相关性模型的图像标注方法, 所述已匹配样本集合通过概率典型相关性分析模型计算 \mathbf{x}'_1 和 \mathbf{x}'_2 。

[0014] 所述的基于弱匹配概率典型相关性模型的图像标注方法, 对应所述未匹配样本集合 $\mathbf{X}_1^{(u)} = \left\{ \left\{ \mathbf{x}'_1 \right\}_{j=N_p+1}^{N_1} \right\}$ 和 $\mathbf{X}_2^{(u)} = \left\{ \left\{ \mathbf{x}'_2 \right\}_{k=N_p+1}^{N_2} \right\}$, \mathbf{x}'_1 和 \mathbf{x}'_2 通过以下公式获得:

$$[0015] \quad P(\mathbf{x}'_1; \theta) = \int P(\mathbf{x}'_1 | \mathbf{z}'_1) P(\mathbf{z}'_1) d\mathbf{z}'_1 \sim N(\mu_1, \mathbf{W}_1 \mathbf{W}_1^T + \psi_1)$$

$$[0016] \quad P(\mathbf{x}'_2; \theta) = \int P(\mathbf{x}'_2 | \mathbf{z}'_2) P(\mathbf{z}'_2) d\mathbf{z}'_2 \sim N(\mu_2, \mathbf{W}_2 \mathbf{W}_2^T + \psi_2)$$

[0017] 其中对于未匹配样本集合 $\mathbf{X}_1^{(u)} = \left\{ \left\{ \mathbf{x}'_1 \right\}_{j=N_p+1}^{N_1} \right\}$ 和 $\mathbf{X}_2^{(u)} = \left\{ \left\{ \mathbf{x}'_2 \right\}_{k=N_p+1}^{N_2} \right\}$, \mathbf{x}'_1 和 \mathbf{x}'_2 则分别由隐变量 \mathbf{z}'_1 和 \mathbf{z}'_2 通过线性变换 \mathbf{W}_1 和 \mathbf{W}_2 附加高斯噪声 ε_1 和 ε_2 获得, ψ_1 , ψ_2 , μ_1 和 μ_2 分别表示随机变量 \mathbf{x}_1 和 \mathbf{x}_2 观察样本集合的协方差和均值。

[0018] 所述的基于弱匹配概率典型相关性模型的图像标注方法, 通过以下公式获取所述弱匹配概率典型相关性模型中成对样本的投影:

$$[0019] \quad E(\mathbf{z}'_1 | \mathbf{x}'_1) = \hat{\mathbf{W}}_1^T (\hat{\mathbf{W}}_1 \hat{\mathbf{W}}_1^T + \hat{\psi}_1)^{-1} (\mathbf{x}'_1 - \hat{\mu}_1)$$

$$[0020] \quad E(\mathbf{z}'_2 | \mathbf{x}'_2) = \hat{\mathbf{W}}_2^T (\hat{\mathbf{W}}_2 \hat{\mathbf{W}}_2^T + \hat{\psi}_2)^{-1} (\mathbf{x}'_2 - \hat{\mu}_2)$$

[0021] 其中, $E(\mathbf{z}'_1 | \mathbf{x}'_1)$, $E(\mathbf{z}'_2 | \mathbf{x}'_2)$ 构成了成对样本 $(\mathbf{x}'_1, \mathbf{x}'_2)$ 在 SemiPCCA 隐空间的典型投影, $\hat{\mathbf{W}}_1$ 和 $\hat{\mathbf{W}}_2$ 表示线性变换矩阵, ε_1 和 ε_2 表示高斯噪声, $\hat{\psi}_1$, $\hat{\psi}_2$, $\hat{\mu}_1$ 和 $\hat{\mu}_2$ 分别表示随机变量 \mathbf{x}_1 和 \mathbf{x}_2 观察样本集合的协方差和均值。

[0022] 本发明还提出一种基于弱匹配概率典型相关性模型的图像标注系统, 包括: 获取已匹配样本集与未匹配样本集模块, 用于获取图像数据库中已标注图像与未标注图像, 分别提取所述已标注图像与所述未标注图像的图像特征和文本特征, 生成已匹配样本集合和未匹配样本集合, 所述已匹配样本集合包括已标注图像特征集合与已标注文本特征集合, 所述未匹配样本集合包括未标注图像特征集合与未标注文本特征集合;

[0023] 训练模型模块,用于根据所述已匹配样本集合与所述未匹配样本集合,训练所述弱匹配概率典型相关性模型;

[0024] 标注模块,用于通过所述弱匹配概率典型相关性模型,对待标注图像进行标注。

[0025] 所述的基于弱匹配概率典型相关性模型的图像标注方法,所述弱匹配概率典型相关性模型的公式为:

$$[0026] \quad \mathbf{D} = \left\{ \left(\mathbf{x}_1^i, \mathbf{x}_2^i \right) \right\}_{i=1}^{N_p} \cup \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1} \cup \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$$

$$[0027] \quad L(\theta) = \prod_{i=1}^{N_p} P(\mathbf{x}_1^i, \mathbf{x}_2^i; \theta) \prod_{j=N_p+1}^{N_1} P(\mathbf{x}_1^j; \theta) \prod_{k=N_p+1}^{N_2} P(\mathbf{x}_2^k; \theta)$$

[0028] 其中, $\mathbf{D} = \left\{ \left(\mathbf{x}_1^i, \mathbf{x}_2^i \right) \right\}_{i=1}^{N_p} \cup \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1} \cup \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$ 表示完整的观察样本集合,包含了匹配和未匹配样本, N_p 为成对观察样本集合的样本数量, $\mathbf{X}_1^{(P)} = \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=1}^{N_1}$ 和 $\mathbf{X}_2^{(P)} = \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=1}^{N_2}$, 其中每一个样本 \mathbf{x}_1^j (\mathbf{x}_2^k) 代表一个 m_1 (m_2) 维向量, $\mathbf{X}_1^{(U)} = \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1}$ 与 $\mathbf{X}_2^{(U)} = \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$ 表示未匹配样本集合, 其中 $\mathbf{X}_1^{(U)}$ 与 $\mathbf{X}_2^{(U)}$ 相互独立生成, 假设样本之间相互独立, 其极大似然值 $L(\theta)$, $P(\mathbf{x}_1^i, \mathbf{x}_2^i; \theta)$ 服从概率典型相关性分析模型, $P(\mathbf{x}_1^j; \theta)$ 和 $P(\mathbf{x}_2^k; \theta)$ 分别表示未匹配样本集合 $\mathbf{X}_1^{(U)} = \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1}$ 和 $\mathbf{X}_2^{(U)} = \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$ 的概率分布。

[0029] 所述的基于弱匹配概率典型相关性模型的图像标注方法,所述已匹配样本集合通过概率典型相关性分析模型计算 \mathbf{x}_1^i 和 \mathbf{x}_2^i 。

[0030] 所述的基于弱匹配概率典型相关性模型的图像标注方法,对应所述未匹配样本集合 $\mathbf{X}_1^{(U)} = \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1}$ 和 $\mathbf{X}_2^{(U)} = \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$, \mathbf{x}_1^j 和 \mathbf{x}_2^k 通过以下公式获得:

$$[0031] \quad P(\mathbf{x}_1^j; \theta) = \int P(\mathbf{x}_1^j | \mathbf{z}_1^j) P(\mathbf{z}_1^j) d\mathbf{z}_1^j \sim N(\mu_1, \mathbf{W}_1 \mathbf{W}_1^T + \psi_1)$$

$$[0032] \quad P(\mathbf{x}_2^k; \theta) = \int P(\mathbf{x}_2^k | \mathbf{z}_2^k) P(\mathbf{z}_2^k) d\mathbf{z}_2^k \sim N(\mu_2, \mathbf{W}_2 \mathbf{W}_2^T + \psi_2)$$

[0033] 其中对于未匹配样本集合 $\mathbf{X}_1^{(U)} = \left\{ \left(\mathbf{x}_1^j \right) \right\}_{j=N_p+1}^{N_1}$ 和 $\mathbf{X}_2^{(U)} = \left\{ \left(\mathbf{x}_2^k \right) \right\}_{k=N_p+1}^{N_2}$, \mathbf{x}_1^j 和 \mathbf{x}_2^k 则分别由隐变量 \mathbf{z}_1^j 和 \mathbf{z}_2^k 通过线性变换 \mathbf{W}_1 和 \mathbf{W}_2 附加高斯噪声 ε_1 和 ε_2 获得, ψ_1 , ψ_2 , μ_1 和 μ_2 分别表示随机变量 x_1 和 x_2 观察样本集合的协方差和均值。

[0034] 所述的基于弱匹配概率典型相关性模型的图像标注方法,通过以下公式获取所述弱匹配概率典型相关性模型中成对样本的投影:

$$[0035] \quad E(\mathbf{z}^i | \mathbf{x}_1^i) = \hat{\mathbf{W}}_1^T (\hat{\mathbf{W}}_1 \hat{\mathbf{W}}_1^T + \hat{\psi}_1)^{-1} (\mathbf{x}_1^i - \hat{\mu}_1)$$

$$[0036] \quad E(\mathbf{z}^i | \mathbf{x}_2^i) = \hat{\mathbf{W}}_2^T (\hat{\mathbf{W}}_2 \hat{\mathbf{W}}_2^T + \hat{\psi}_2)^{-1} (\mathbf{x}_2^i - \hat{\mu}_2)$$

[0037] 其中, $E(\mathbf{z}^i | \mathbf{x}_1^i)$, $E(\mathbf{z}^i | \mathbf{x}_2^i)$ 构成了成对样本 $(\mathbf{x}_1^i, \mathbf{x}_2^i)$ 在 SemiPCCA 隐空间的典型投影, $\hat{\mathbf{W}}_1$ 和 $\hat{\mathbf{W}}_2$ 表示线性变换矩阵, ε_1 和 ε_2 表示高斯噪声, $\hat{\psi}_1$, $\hat{\psi}_2$, $\hat{\mu}_1$ 和 $\hat{\mu}_2$ 分别表示随机变量 x_1 和 x_2 观察样本集合的协方差和均值。

[0038] 由以上方案可知,本发明的优点在于:

[0039] 本发明关注于各模态内部的全局结构,模型参数的估计受到了未匹配样本的影响,而未匹配样本则揭示了各模态样本空间的全局结构;

[0040] 在人工弱匹配多模态数据集上的实验表明,本发明可以有效克服传统 CCA 和 PCCA 在匹配样本不足的情况下出现的过拟合问题,取得了很好的效果;

[0041] 基于本发明的图像自动标注方法,同时使用标注图像及其关键词和未标注图像学习视觉模态和文本模态之间的关联,从而能更准确地对未知图像进行标注。

附图说明

[0042] 图 1 为本发明整体流程图;

[0043] 图 2 为 SemiPCCA 的图模型;

[0044] 图 3 为加权余弦距离图;

[0045] 图 4a 为 CCA、PCCA 和 SemiPCCA 获得的典型投影向量图;

[0046] 图 4b 为另一 CCA、PCCA 和 SemiPCCA 获得的典型投影向量图;

[0047] 图 5a 为图像语义标注方法图;

[0048] 图 5b 为另一图像语义标注方法图;

[0049] 图 5c 为又一图像语义标注方法图;

[0050] 图 6 为在 PCCA 隐空间中未标注图像和未标注图像的后验概率分布图;

[0051] 图 7 为 Core15K 标注图像和未标注图像在 PCCA 隐空间的后验概率分布图;

[0052] 图 8 为 Core1 图像库示例图;

[0053] 图 9 为 0 阶 Color-HLAC 特征和 1 阶 Color-HLAC 特征模板。

[0054] 其中附图标记为:

[0055] 步骤 101/102/103。

具体实施方式

[0056] 以下为本发明方法的具体步骤,如图 1 所示:

[0057] 步骤 101,获取图像数据库中已标注图像与未标注图像,分别提取所述已标注图像与所述未标注图像的图像特征和文本特征,生成已匹配样本集合和未匹配样本集合,所述已匹配样本集合包括已标注图像特征集合与已标注文本特征集合,所述未匹配样本集合包括未标注图像特征集合与未标注文本特征集合;

[0058] 步骤 102,根据所述已匹配样本集合与所述未匹配样本集合,训练所述弱匹配概率典型相关性模型;

[0059] 步骤 103,通过所述弱匹配概率典型相关性模型,对待标注图像进行标注。

[0060] 以下为本发明中训练弱匹配概率典型相关性模型的具体步骤,如下所示:

[0061] 给定数量为 N_p 的成对观察样本集合 $\mathbf{X}_1^{(p)} = \{(\mathbf{x}_1^i)\}_{i=1}^{N_p}$ 和 $\mathbf{X}_2^{(p)} = \{(\mathbf{x}_2^i)\}_{i=1}^{N_p}$, 其中每一个样本 \mathbf{x}_1^i (\mathbf{x}_2^i) 代表一个 m_1 (m_2) 维向量。在成对样本数量很小的情况下,CCA 建立的相关性模型容易出现过拟合问题。下面,本发明考虑给出未匹配样本集合 $\mathbf{X}_1^{(u)} = \{(\mathbf{x}_1^i)\}_{i=N_p+1}^{M_1}$ 与 / 或

$\mathbf{X}_2^{(U)} = \left\{ \left\{ \mathbf{x}_2^k \right\}_{k=N_p+1}^{N_2} \right\}$, 其中 $\mathbf{X}_1^{(U)}$ 与 $\mathbf{X}_2^{(U)}$ 相互独立生成, 为了解决传统 CCA 和概率典型相关性分析 (PCCA) 模型无法直接处理未匹配样本的弊端, 本发明提出一种全新的弱匹配概率典型相关性模型 (Semi-paired PCCA, 简称 SemiPCCA)。SemiPCCA 充分利用未匹配样本解决过拟合问题, 图 2 给出了 SemiPCCA 的图模型。

[0062] $\mathbf{D} = \left\{ \left\{ \mathbf{x}_1^i, \mathbf{x}_2^i \right\}_{i=1}^{N_p} \cup \left\{ \left\{ \mathbf{x}_1^j \right\}_{j=N_p+1}^{N_1} \cup \left\{ \left\{ \mathbf{x}_2^k \right\}_{k=N_p+1}^{N_2} \right\} \right\}$ 表示完整的观察样本集合, 包含了匹配和未匹配样本。假设样本之间项目独立, 其极大似然值如下:

$$[0063] \quad L(\theta) = \prod_{i=1}^{N_p} P(\mathbf{x}_1^i, \mathbf{x}_2^i; \theta) \prod_{j=N_p+1}^{N_1} P(\mathbf{x}_1^j; \theta) \prod_{k=N_p+1}^{N_2} P(\mathbf{x}_2^k; \theta)$$

[0064] 在 SemiPCCA 模型中, 对于成对样本 $\left\{ \left\{ \mathbf{x}_1^i, \mathbf{x}_2^i \right\}_{i=1}^{N_p} \right\}$, \mathbf{x}_1^i 和 \mathbf{x}_2^i 由相同的隐变量 z^i 生成, 且 $P(\mathbf{x}_1^i, \mathbf{x}_2^i; \theta)$ 服从概率典型相关性分析 (PCCA) 模型, 即

$$[0065] \quad P(\mathbf{x}_1^i, \mathbf{x}_2^i; \theta) \sim N \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \mathbf{W}_1 \mathbf{W}_1^T + \psi_1 & \mathbf{W}_1 \mathbf{W}_2^T \\ \mathbf{W}_2 \mathbf{W}_1^T & \mathbf{W}_2 \mathbf{W}_2^T + \psi_2 \end{pmatrix} \right)$$

[0066] 对于未匹配样本集合 $\mathbf{X}_1^{(U)} = \left\{ \left\{ \mathbf{x}_1^j \right\}_{j=N_p+1}^{N_1} \right\}$ 和 $\mathbf{X}_2^{(U)} = \left\{ \left\{ \mathbf{x}_2^k \right\}_{k=N_p+1}^{N_2} \right\}$, \mathbf{x}_1^j 和 \mathbf{x}_2^k 则分别由隐变量 z_1^j 和 z_2^k 通过线性变换 W_1 和 W_2 附加高斯噪声 ε_1 和 ε_2 获得, 即

$$[0067] \quad P(\mathbf{x}_1^j; \theta) = \int P(\mathbf{x}_1^j | z_1^j) P(z_1^j) dz_1^j \sim N(\mu_1, \mathbf{W}_1 \mathbf{W}_1^T + \psi_1)$$

$$[0068] \quad P(\mathbf{x}_2^k; \theta) = \int P(\mathbf{x}_2^k | z_2^k) P(z_2^k) dz_2^k \sim N(\mu_2, \mathbf{W}_2 \mathbf{W}_2^T + \psi_2)$$

[0069] SemiPCCA 模型中, 成对样本 $(\mathbf{x}_1^i, \mathbf{x}_2^i)$ 的投影方法类似 PCCA 模型, 即

$$[0070] \quad E(z^i | \mathbf{x}_1^i) = \hat{W}_1^T (\hat{W}_1 \hat{W}_1^T + \hat{\psi}_1)^{-1} (\mathbf{x}_1^i - \hat{\mu}_1)$$

$$[0071] \quad E(z^i | \mathbf{x}_2^i) = \hat{W}_2^T (\hat{W}_2 \hat{W}_2^T + \hat{\psi}_2)^{-1} (\mathbf{x}_2^i - \hat{\mu}_2)$$

[0072] $E(z^i | \mathbf{x}_1^i)$, $E(z^i | \mathbf{x}_2^i)$ 构成了样本空间到 SemiPCCA 隐空间的典型投影。虽然 SemiPCCA 模型投影的结果看似和 PCCA 模型相同, 但 \hat{W}_1 和 \hat{W}_2 的计算却受到了未匹配样本的影响, 而未匹配样本则揭示了各领域样本空间的全局结构。同时为了使相关度最大化, 不同样本空间的投影向量之间也会相互影响。

[0073] EM 算法求解 SemiPCCA, 如下所示:

[0074] 考虑到观察样本的极大似然函数 $L(\theta)$ 由三部分构成, 因此 E 步骤, 我们需要分别处理。

[0075] 对于匹配样本集合 $\left\{ \left\{ \mathbf{x}_1^i, \mathbf{x}_2^i \right\}_{i=1}^{N_p} \right\}$ 中的第 i 对样本 $(\mathbf{x}_1^i, \mathbf{x}_2^i)$, 我们给出隐变量 z^i 的后验概率, 即:

$$[0076] \quad P(z^i | \mathbf{x}_1^i, \mathbf{x}_2^i; \theta) \sim N \left(\mathbf{W}^T (\mathbf{W} \mathbf{W}^T + \psi)^{-1} \begin{pmatrix} \mathbf{x}_1^i \\ \mathbf{x}_2^i \end{pmatrix} - \mu, \mathbf{I} - \mathbf{W}^T (\mathbf{W} \mathbf{W}^T + \psi)^{-1} \mathbf{W} \right)$$

[0077] 根据该后验概率 $P(z^i | \mathbf{x}_1^i, \mathbf{x}_2^i; \theta)$, 我们计算得到 z^i 和 $z^i z^{iT}$ 的期望值:

$$[0078] \quad \langle z^i \rangle = W^T (WW^T + \psi)^{-1} \left(\begin{pmatrix} x_1^i \\ x_2^i \end{pmatrix} - \mu \right)$$

$$[0079] \quad \langle z^i z^{iT} \rangle = \langle z^i \rangle \langle z^i \rangle^T + I - W^T (WW^T + \psi)^{-1} W$$

[0080] 对于未匹配样本 $\{(x_j^i)\}_{j=N_p+1}^{N_1}$, 隐变量 z_j^i 只受 x_j^i 影响, 其后验概率的计算如下所示:

$$[0081] \quad P(z_j^i | x_j^i; \theta) \sim N \left(W_1^T (W_1 W_1^T + \psi_1)^{-1} (x_j^i - \mu_1), I - W_1^T (W_1 W_1^T + \psi_1)^{-1} W_1 \right)$$

[0082] 根据该后验概率 $P(z_j^i | x_j^i; \theta)$, 我们计算得到 z_j^i 和 $z_j^i z_j^{iT}$ 的期望值:

$$[0083] \quad \langle z_j^i \rangle = W_1^T (W_1 W_1^T + \psi_1)^{-1} (x_j^i - \mu_1)$$

$$[0084] \quad \langle z_j^i z_j^{iT} \rangle = \langle z_j^i \rangle \langle z_j^i \rangle^T + I - W_1^T (W_1 W_1^T + \psi_1)^{-1} W_1$$

[0085] 对于未匹配样本 $\{(x_k^k)\}_{k=N_p+1}^{N_2}$, 隐变量 z_k^k 只受 x_k^k 影响, 其后验概率的计算如下所示:

$$[0086] \quad P(z_k^k | x_k^k; \theta) \sim N \left(W_2^T (W_2 W_2^T + \psi_2)^{-1} (x_k^k - \mu_2), I - W_2^T (W_2 W_2^T + \psi_2)^{-1} W_2 \right)$$

[0087] 根据该后验概率 $P(z_k^k | x_k^k; \theta)$, 我们计算得到 z_k^k 和 $z_k^k z_k^{kT}$ 的期望值:

$$[0088] \quad \langle z_k^k \rangle = W_2^T (W_2 W_2^T + \psi_2)^{-1} (x_k^k - \mu_2)$$

$$[0089] \quad \langle z_k^k z_k^{kT} \rangle = \langle z_k^k \rangle \langle z_k^k \rangle^T + I - W_2^T (W_2 W_2^T + \psi_2)^{-1} W_2$$

[0090] M 步骤, 固定 E 步骤计算得到的 $P(z^i | x_1^i, x_2^i; \theta)$, $P(z_j^i | x_j^i; \theta)$ 和 $P(z_k^k | x_k^k; \theta)$, 通过偏导数计算似然 $L(\theta)$ 最大化时, 相应参数的取值。

[0091] 对于 x_1 和 x_2 的均值,

$$[0092] \quad \tilde{\mu}_1 = \hat{\mu}_1 = \frac{1}{N_1} \sum_{i=1}^{N_1} x_1^i, \quad \tilde{\mu}_2 = \hat{\mu}_2 = \frac{1}{N_2} \sum_{k=1}^{N_2} x_2^k$$

[0093] 由于 EM 算法迭代过程中, $\tilde{\mu}_1$ 和 $\tilde{\mu}_2$ 的取值不变, 我们可以通过中心化样本集合 $X_1^{(p)} \cup X_1^{(u)}$, $X_2^{(p)} \cup X_2^{(u)}$, 避免学习过程中重复学习。为了简化描述, 下文中 x_1^i , x_2^k , x_1^i 和 x_2^k 均表示经过中心化的向量。

[0094] 对于投影向量集合, 我们获得以下更新公式:

$$[0095] \quad \tilde{W}_1 = \left[\sum_{i=1}^{N_p} x_1^i \langle z^i \rangle^T + \sum_{j=N_p+1}^{N_1} x_1^j \langle z_j^i \rangle^T \right] \left[\sum_{i=1}^{N_p} \langle z^i z^{iT} \rangle + \sum_{j=N_p+1}^{N_1} \langle z_j^i z_j^{iT} \rangle \right]^{-1}$$

$$[0096] \quad \tilde{W}_2 = \left[\sum_{k=1}^{N_p} x_2^k \langle z^k \rangle^T + \sum_{k=N_p+1}^{N_2} x_2^k \langle z_k^k \rangle^T \right] \left[\sum_{k=1}^{N_p} \langle z^k z^{kT} \rangle + \sum_{k=N_p+1}^{N_2} \langle z_k^k z_k^{kT} \rangle \right]^{-1}$$

[0097] 对于高斯噪声的方差, 我们获得以下更新公式:

$$[0098] \quad \tilde{\psi}_1 = \frac{1}{N_1} \left\{ \sum_{i=1}^{N_p} (x_1^i - \tilde{W}_1 \langle z^i \rangle) (x_1^i - \tilde{W}_1 \langle z^i \rangle)^T + \sum_{j=N_p+1}^{N_1} (x_1^j - \tilde{W}_1 \langle z_j^i \rangle) (x_1^j - \tilde{W}_1 \langle z_j^i \rangle)^T \right\}$$

$$[0099] \quad \tilde{\psi}_2 = \frac{1}{N_2} \left\{ \sum_{k=1}^{N_p} (x_2^k - \tilde{W}_2 \langle z^k \rangle) (x_2^k - \tilde{W}_2 \langle z^k \rangle)^T + \sum_{k=N_p+1}^{N_2} (x_2^k - \tilde{W}_2 \langle z_k^k \rangle) (x_2^k - \tilde{W}_2 \langle z_k^k \rangle)^T \right\}$$

[0100] 以下为本发明具体实施例,如下所示:

[0101] 人工弱匹配多模态数据集上的实验,如下所示:

[0102] 为了验证 SemiPCCA 模型的有效性,我们构造以下人工数据集:样本集合 $\{z_i\}_{i=1}^N$ 服从 $N(0, I_d)$, 其中维度 $d = 2$, 样本数量 $N = 300$, 完整的匹配样本集合 $\{(x_1^i, x_2^i)\}_{i=1}^N$ 通过以下方式构造获得,

$$[0103] \quad x_1 = T_1 z + \varepsilon_1, T_1 \in \mathbb{R}^{m_1 \times d}$$

$$[0104] \quad x_2 = T_2 z + \varepsilon_2, T_2 \in \mathbb{R}^{m_2 \times d}$$

$$[0105] \quad \text{其中, } P(\varepsilon_1) \sim N\left(0, \begin{bmatrix} 0.75 & 0.5 \\ 0.5 & 0.75 \end{bmatrix}\right), \quad P(\varepsilon_2) \sim N\left(0, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right), \quad T_1 = \begin{bmatrix} 0.6 & -1/\sqrt{2} \\ 0.8 & -1/\sqrt{2} \end{bmatrix}, \quad T_2 = \begin{bmatrix} 0.3 & -0.7 \\ 0.4 & 0.7 \end{bmatrix},$$

样本维度分别设置为 $m_1 = 2, m_2 = 2$;

[0106] 为了获得弱匹配的样本集合,我们构造一个判别函数 $f(x_2) = a^T x_2 - \theta$, 其中 $a = (a_1, \dots, a_{m_2})^T$, θ 表示判别阈值,对于样本 (x_1^i, x_2^i) , 如果其判别函数值 $f(x_2^i) < 0$, 则从 $\{x_2^i\}_{i=1}^N$ 中移除样本。可见, θ 越大, 移除的样本就越多;

[0107] 在比较 SemiPCCA 与传统 CCA 和 PCCA 时,我们选择了以下加权余弦距离,

$$[0108] \quad c(W_x, W_x^*, A^*) = \sum_{i=1}^d \lambda_i^* \frac{w_{x,i}^T w_{x,i}^*}{\|w_{x,i}\| \|w_{x,i}^*\|}$$

[0109] 其中, $W_x^* = (w_{x,1}^*, w_{x,2}^*, \dots, w_{x,d}^*)^T$ 和 $A^* = \text{diag}(\lambda_1^*, \lambda_2^*, \dots, \lambda_d^*)$ 分别表示有完整匹配样本通过 CCA 分析后,获得的“真正” d 个典型投影向量和相关系数;

[0110] 图 3 给出了判别阈值 θ 在 -2 到 5 的取值范围内,经过 1000 次独立实验获得的加权余弦距离平均值。实验结果表明随着判别阈值 θ 的提高,匹配样本逐渐减少, SemiPCCA 模型由于考虑了为匹配样本,其性能明显好于传统 CCA 和 PCCA,解决了过拟合问题;

[0111] 图 4(a) (图 4(b)) 描述了 $\theta = -2$ ($\theta = 4$) 时,匹配样本 (蓝色方形)、未匹配样本 (红色圆形) 的分布情况,以及分别由 CCA、PCCA 和 SemiPCCA 获得的典型投影向量。在只考虑了匹配样本的情况下,CCA 和 PCCA 出现了过拟合问题。

[0112] 以下为图像语义标注,如下所示:

[0113] 图像检索技术包括两种主流解决方案:基于文本的图像检索和基于内容的图像检索。基于文本的图像检索利用人工对图像进行标注,并在此基础上利用传统的文本搜索引擎查询图像,这种查询方式比较直观,但是,人工标注费时费力,使得这种检索技术不能推广到大规模的图像数据库。基于内容的图像检索采用特征提取和高维索引技术进行图像检索,它为每幅图像提取底层视觉特征,以高维形式存入数据库,通过比较这些特征的相似度来获得检索结果。这种技术在人脸识别、商标识别等某些特殊领域得到了很好的应用,但由于存在语义鸿沟,视觉特征相似的图像很可能在语义上是不相关的。为了获得语义相关的检索结果,同时避免大量的手工标注,图像自动标注成为当前关键的具有挑战性的课题;

[0114] 图像标注方法可分为有监督的分类算法和关联建模。有监督的分类算法是一种最直接的图像标注方法。有监督的分类算法将各个语义类别 (一个关键词或关键词集合) 看作独立的概念,通过训练一组经过语义标注的样本图像,为每个语义类别建立各不相同

的二类分类器,然后利用分类器将未标注或未归类的图像归并到某一语义类,如图 5(a) 所示;

[0115] 关联建模的方法从文本领域的研究得到启发。这类方法利用现有的标注好的图像数据集,在无监督的基础上学习图像的视觉特征和文本关键词之间的关联,然后通过统计推理的方法将这种关联应用到未标注的图像。关联建模的基本思想是引入随机变量对客观世界的隐藏状态 L 进行编码,随机变量的各个状态定义了语义关键词和图像特征的联合分布。不同的标注方法对于隐藏状态做了不同的定义。有些方法将图像或图像聚类与隐藏状态联系,如图 5(b) 所示,还有方法同时使用图像和关键词估计隐藏变量的分布,实现某些模型的高层次分组(如主题)与隐藏状态相联系,如图 5(c) 所示。

[0116] 以下为学习与标注,如下所示:

[0117] 类似图 5(c) 所示的关联建模方法,本发明提出了基于 SemiPCCA 的图像标注方法:

[0118] 对于已标注样本 (x_1, x_2) ,隐空间中随机变量 z 的后验概率 $P(z | x_1, x_2)$ 服从以下均值 \hat{z}_{12} 、方差 Ψ_{12} 高斯分布:

$$[0119] \quad \hat{z}_{12} = E(z | x_1, x_2) = \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}^T \begin{pmatrix} (I - P_d^2)^{-1} & -(I - P_d^2)^{-1} P_d \\ -(I - P_d^2)^{-1} P_d & (I - P_d^2)^{-1} \end{pmatrix} \begin{pmatrix} U_{1d}^T (x_1 - \hat{\mu}_1) \\ U_{2d}^T (x_2 - \hat{\mu}_2) \end{pmatrix}$$

$$[0120] \quad \Psi_{12} = \text{var}(z | x_1, x_2) = I - \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}^T \begin{pmatrix} (I - P_d^2)^{-1} & -(I - P_d^2)^{-1} P_d \\ -(I - P_d^2)^{-1} P_d & (I - P_d^2)^{-1} \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}$$

[0121] 类似地,对于未标注的样本,隐空间中随机变量 z 在只给定样本图像特征 x_1 的情况下,其后验概率 $P(z | x_1)$ 服从以下均值 \hat{z}_1 、方差 Ψ_1 高斯分布:

[0122]

$$\hat{z}_1 = E(z | x_1) = M_1^T U_{1d}^T (x_1 - \hat{\mu}_1)$$

$$[0123] \quad \Psi_1 = \text{var}(z | x_1) = I - M_1 M_1^T$$

[0124] 根据上述结论,对于已标注图像和未标注图像,隐空间中随机变量 z 的分布情况如图 6 所示,图 7 给出了来自 Core15K 图像库中 100000 (Bears) 和 109000 (Foxes&Coyotes) 两个主题的 6 幅标注图像和 2 幅未标注图像在隐空间中的后验概率分布;

[0125] 设 $\{T_i^{(p)} = (x_1^i, x_2^i)\}_{i=1}^{N_p}$ 表示已标注图像的图像特征和文本关键词集合, $\{Q_j^{(u)} = (x_1^j)\}_{j=N_p+1}^N$ 表示未标注图像的图像特征集合。Nakayam 等人提出了一种基于 PCCA 的图像标注方法。对于给定未标注图像 $Q_j^{(u)}$,文本关键词的后验概率如式 (18):

$$[0126] \quad P(w | Q_j^{(u)}) = \sum_{i=1}^{N_p} P(w | T_i^{(p)}) P(T_i^{(p)} | Q_j^{(u)}), \quad (1)$$

[0127] 其中,

$$[0128] \quad P(T_i^{(p)} | Q_j^{(u)}) = \frac{D_{\text{KL}}(P(z | T_i^{(p)}), P(z | Q_j^{(u)}))}{\sum_{k=1}^{N_p} D_{\text{KL}}(P(z | T_k^{(p)}), P(z | Q_j^{(u)}))},$$

[0129] $D_{KL}\left(\mathbf{P}\left(\mathbf{z}|\mathbf{T}_i^{(p)}\right), \mathbf{P}\left(\mathbf{z}|\mathbf{Q}_j^{(u)}\right)\right)$ 表示分布 $\mathbf{P}\left(\mathbf{z}|\mathbf{T}_i^{(p)}\right)$ 和 $\mathbf{P}\left(\mathbf{z}|\mathbf{Q}_j^{(u)}\right)$ 在隐空间中的 KL 距离之和, 即:

$$[0130] \quad D_{KL}\left(\mathbf{P}\left(\mathbf{z}|\mathbf{T}_i^{(p)}\right), \mathbf{P}\left(\mathbf{z}|\mathbf{Q}_j^{(u)}\right)\right) = KL\left(\mathbf{P}\left(\mathbf{z}|\mathbf{T}_i^{(p)}\right), \mathbf{P}\left(\mathbf{z}|\mathbf{Q}_j^{(u)}\right)\right) + KL\left(\mathbf{P}\left(\mathbf{z}|\mathbf{Q}_j^{(u)}\right), \mathbf{P}\left(\mathbf{z}|\mathbf{T}_i^{(p)}\right)\right);$$

[0131] 根据多维高斯分布间 KL 距离的计算公式, 得:

$$[0132] \quad KL\left(\mathbf{P}\left(\mathbf{z}|\mathbf{T}_i^{(p)}\right), \mathbf{P}\left(\mathbf{z}|\mathbf{Q}_j^{(u)}\right)\right) = \frac{1}{2} \left[\text{tr}\left(\boldsymbol{\psi}_i^{-1} \boldsymbol{\psi}_{j2}\right) - d - \log\left(\frac{|\boldsymbol{\psi}_i|}{|\boldsymbol{\psi}_{j2}|}\right) + \left(\hat{\mathbf{z}}_q - \hat{\mathbf{z}}_i\right)^T \boldsymbol{\psi}_i^{-1} \left(\hat{\mathbf{z}}_q - \hat{\mathbf{z}}_i\right) \right].$$

[0133] 对于不同的样本, 上式的前三项是常数, 所以可以将以上 KL 距离简化为:

$$[0134] \quad KL\left(\mathbf{P}\left(\mathbf{z}|\mathbf{T}_i^{(p)}\right), \mathbf{P}\left(\mathbf{z}|\mathbf{Q}_j^{(u)}\right)\right) = \frac{1}{2} \left(\hat{\mathbf{z}}_q - \hat{\mathbf{z}}_i\right)^T \boldsymbol{\psi}_i^{-1} \left(\hat{\mathbf{z}}_q - \hat{\mathbf{z}}_i\right), \quad (2)$$

[0135] 同理,

$$[0136] \quad KL\left(\mathbf{P}\left(\mathbf{z}|\mathbf{Q}_j^{(u)}\right), \mathbf{P}\left(\mathbf{z}|\mathbf{T}_i^{(p)}\right)\right) = \frac{1}{2} \left(\hat{\mathbf{z}}_q - \hat{\mathbf{z}}_i\right)^T \boldsymbol{\psi}_{j2}^{-1} \left(\hat{\mathbf{z}}_q - \hat{\mathbf{z}}_i\right). \quad (3)$$

[0137] $\mathbf{P}\left(w|\mathbf{T}_i^{(p)}\right)$ 定义如下:

$$[0138] \quad \mathbf{P}\left(w|\mathbf{T}_i^{(p)}\right) = \mu \delta_{w, \mathbf{T}_i^{(p)}} + (1 - \mu) \frac{N_w}{NW}, \quad (4)$$

[0139] 其中, N_w 表示标注图像集合中包含语义关键字 w 的图像数量, NW 表示语义关键字的数量, $\delta_{w, \mathbf{T}_i^{(p)}} = 1$ 表示标注样本 $\mathbf{T}_i^{(p)}$ 包含语义关键字 w , 否则 $\delta_{w, \mathbf{T}_i^{(p)}} = 0$, 参数 $0 < \mu < 1$ (取 $\mu = 0.99$)。

[0140] 以下为本发明的实验过程和结果, 如下所示:

[0141] 实验数据:

[0142] 图像标注领域, 很多研究者采用 Corel 图像库进行测试, 图 8 给出了 Corel 图像库中的一些图像示例。使用 Corel 图像库所得到的结果在一定程度上能为判断标注系统性能提供依据。

[0143] 我们采用 Corel5K 数据集和 Corel30K 数据集进行实验。Corel5K 数据集包含 5000 幅图像, 来自 50 个 Corel 库存图像 CD, 每张 CD 包含同样语义内容的 100 幅图像, 每幅图像标注 1~5 个关键词。Corel5k 共有 371 个关键词, 将至少标注了 8 幅图像的关键词选入词汇表, 合计 260 个关键词。整个数据集分为三部分: 4500 幅标注图像作为训练集, 500 幅图像作为测试集。Corel30K 数据集与 Corel5K 类似, 但包含 31,695 幅图像和 5,587 个关键词, 将至少标注了 10 幅图像的关键词选入词汇表, 合计 950 个关键词。

[0144] 我们使用 Corel5K 测试集的 500 幅图像作为测试图像, 从 Corel5K 训练集中的分别选择 1500 幅、2250 幅和 4500 幅图像作为标注图像, 其余图像与 Corel30K 数据集集中的 31,695 幅图像作为未标注图像, 参与 SemiPCCA 模型的学习。

[0145] 以下为本发明获取图像特征, 如下所示:

[0146] 由于无法确定图像中对象的位置和数量, 所以通常的做法是基于分块、基于特征点和基于对象 3 种方法划分图像, 抽取各个区域的底层特征, 并进行聚类, 每个类心看作一个“视觉词”, 所有“视觉词”形成的全集称为“视觉词汇表”, 图像表示为“视觉词袋”, 然而, 不正确的图像划分会导致算法识别性能的降低。

[0147] 为了实现无分割的图像标注, 本发明使用基于图像颜色的高阶局部自相关特征

(Color Higher order Local Auto-Correlation, 简称 Color-HLAC), 1988 年, 东京大学教授 Nobuyuki Otsu 提出了高阶局部自相关 (Higher order Local Auto-Correlation, 简称 HLAC) 的概念。HLAC 使用模板匹配的方法快速计算二值图像相邻像素点的自相关特征, 能够很好地提取图像的局部信息, 描述空间上的相关关系, m -th HLAC 表示 m 阶 HLAC 特征, 随着阶数的增加, HLAC 特征的代表能力增强, 但同时计算量也在增加, 所以通常使用 1 阶 HLAC 特征 (1st HLAC) 或 2 阶 HLAC 特征 (2nd HLAC 特征), HLAC 已经被广泛地用于图像识别。Color-HLAC 特征是 HLAC 特征在 RGB 图像上的扩展, 分别计算 RGB 各层的 HLAC 特征, 然后“串行融合”, 1 阶 Color-HLAC 特征为 45 维, 模板如图 9 所示。2 阶 Color-HLAC 特征为 714 维。

[0148] 以下为本发明图像自动标注结果, 如下所示:

[0149] 图像标注的性能通过比较测试集的图像自动标注结果与原始标注进行评价。本实验取前 5 个后验概率最大的关键词作为每幅图像的标注结果, 并计算测试集中每个关键字的精度 (也称查准率)、召回率 (也称查全率) 及其综合评价指标 F1 值。对于一个关键词 w , 精度 $Precision = B/A$, 召回率 $Recall = B/C$, 综合评价指标 $F1 = 2 * P * R / (P + R)$, 其中 A 表示所有自动标注了 w 的图像个数, B 表示正确标注 w 的图像个数, 即这些图像的原始标注和自动标注都包含 w , C 表示原始标注中包含 w 的图像个数。计算精度和召回率的平均值可以用来评价系统的标注性能。此外, 本实验也考虑了召回率大于 0 的关键词个数, 这个值可以代表系统能够有效学习的关键词个数。

[0150] 使用平均精度和平均召回率比较了若干图像自动标注方法的性能。表 1 给出了 PCCA 和 SemiPCCA 在 Core1 图像库的标注性能比较, 包括性能最佳的 49 个关键词的平均召回率和平均精度, 以及全部 260 个关键词的平均召回率和平均精度, 训练集分别选择 Core15k 中的 1500 幅、2250 幅和 4500 幅标注图像, 表 1 中数据可以看出, SemiPCCA 的性能大大优于 PCCA。

[0151] 表 1 PCCA, SemiPCCA 在 Core1 图像库上的图像自动标注性能比较

[0152]

| 模型 | PCCA | | | SemiPCCA | | |
|--------------------------|------|------|------|----------|------|------|
| 训练样本 | 1500 | 2250 | 4500 | 1500 | 2250 | 4500 |
| 召回率>0 的关键词个数 | 99 | 126 | 150 | 113 | 132 | 151 |
| 性能最佳的 49 个关键词的平均召回率和平均精度 | | | | | | |
| 平均召回率 | 0.61 | 0.74 | 0.89 | 0.71 | 0.85 | 0.94 |
| 平均精度 | 0.56 | 0.65 | 0.7 | 0.60 | 0.72 | 0.77 |
| F1 值 | 0.58 | 0.69 | 0.78 | 0.65 | 0.78 | 0.85 |
| 全部 260 个关键词的平均召回率和平均精度 | | | | | | |
| 平均召回率 | 0.16 | 0.24 | 0.30 | 0.20 | 0.27 | 0.32 |
| 平均精度 | 0.13 | 0.18 | 0.22 | 0.15 | 0.20 | 0.24 |
| F1 值 | 0.14 | 0.21 | 0.25 | 0.17 | 0.23 | 0.27 |

[0153] 表 2 给出了 TM, CMRM, CRM, MBRM, PLSA-WORDS, GM-PLSA 和本文提出的 SemiPCCA 的标注性能。为了与过去的模型进行比较, 训练集采用 Core15k 中 4500 幅标注图像, 表中同样报告了两种标注结果: 性能最佳的 49 个关键词的平均召回率和平均精度与全部 260 个关键词的平均召回率和平均精度, 表 2 中数据可以看出, SemiPCCA 的性能大大优于 TM, CMRM, CRM 和 PLSA-WORDS, 也稍优于 MBRM 和 GM-PLSA。

[0154] 表 2 SemiPCCA 与其他模型在 Core15k 图像库上的图像自动标注性能比较

[0155]

| 模型 | TM | CMRM | CRM | MBRM | PLSA-WO RDS | GM-PLSA | Semi-PCCA |
|--------------------------|------|------|------|------|----------------|---------|-------------|
| 召回率>0 的关键词个数 | 49 | 66 | 107 | 122 | 105 | 125 | 151 |
| 性能最佳的 49 个关键词的平均召回率和平均精度 | | | | | | | |
| 平均召回率 | 0.34 | 0.48 | 0.70 | 0.78 | 0.71 | 0.79 | 0.94 |
| 平均精度 | 0.20 | 0.40 | 0.59 | 0.74 | 0.56 | 0.76 | 0.77 |
| F1 值 | 0.25 | 0.44 | 0.64 | 0.76 | 0.63 | 0.77 | 0.85 |
| 全部 260 个关键词的平均召回率和平均精度 | | | | | | | |
| 平均召回率 | 0.04 | 0.09 | 0.19 | 0.25 | 0.20 | 0.25 | 0.32 |
| 平均精度 | 0.06 | 0.10 | 0.24 | 0.24 | 0.14 | 0.26 | 0.24 |
| F1 值 | 0.05 | 0.09 | 0.21 | 0.24 | 0.16 | 0.25 | 0.27 |

[0156] 针对弱匹配多模态数据的相关性建模问题,本发明提出了一种全新的弱匹配概率典型相关性模型 (SemiPCCA),不同于以往的弱匹配典型相关性分析模型, SemiPCCA 完全基于概率典型相关性分析模型 (PCCA),关注于各模态内部的全局结构,模型参数的估计受到了未匹配样本的影响,而未匹配样本则揭示了各模态样本空间的全局结构。在人工弱匹配多模态数据集上的实验表明, SemiPCCA 可以有效克服传统 CCA 和 PCCA 在匹配样本不足的情况下出现的过拟合问题,取得了很好的效果。

[0157] 接着,本发明提出了一种基于 SemiPCCA 的图像自动标注方法,该方法是基于关联建模的思想,同时使用标注图像及其关键词和未标注图像估计隐空间的分布,学习视觉模态和文本模态之间的关联,从而能较好地对未知图像进行标注。在 Core1 数据集上进行的实验表明, SemiPCCA 比几种典型的图像标注方法具有更高的标注精度和更好的检索效果。

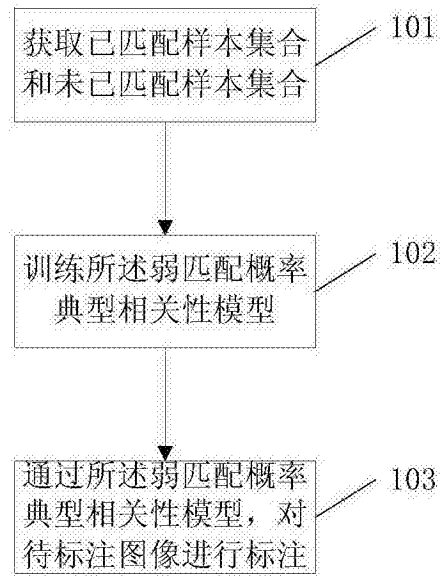


图 1

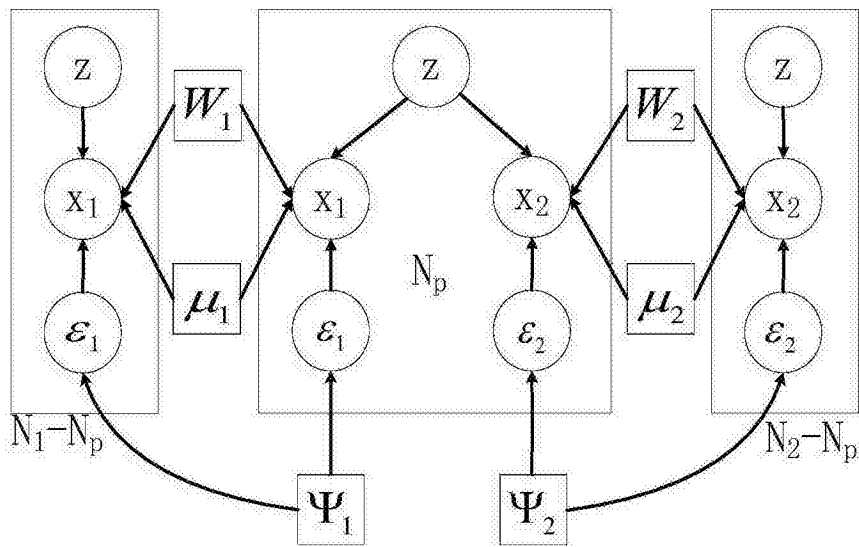


图 2

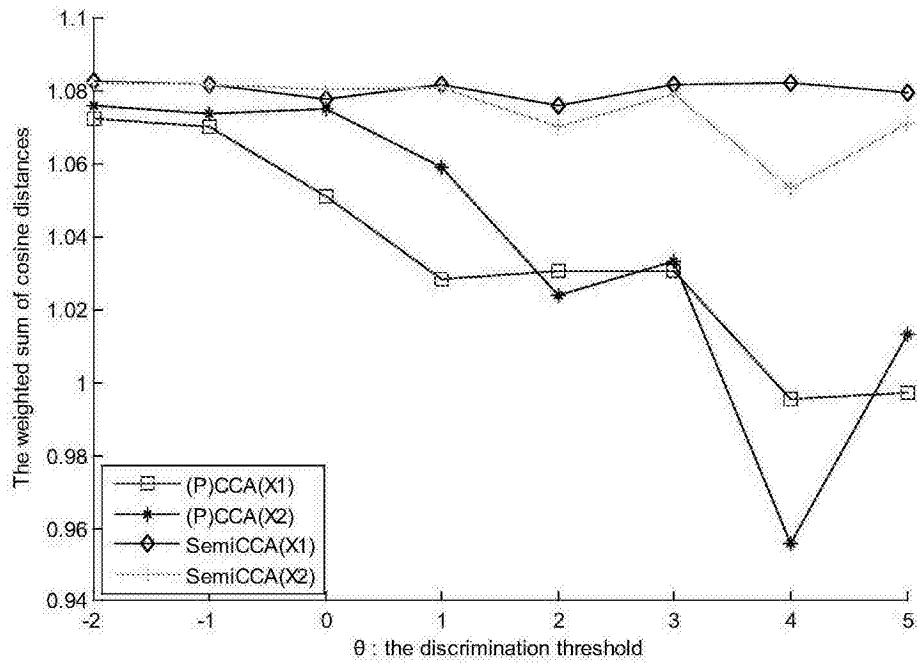


图 3

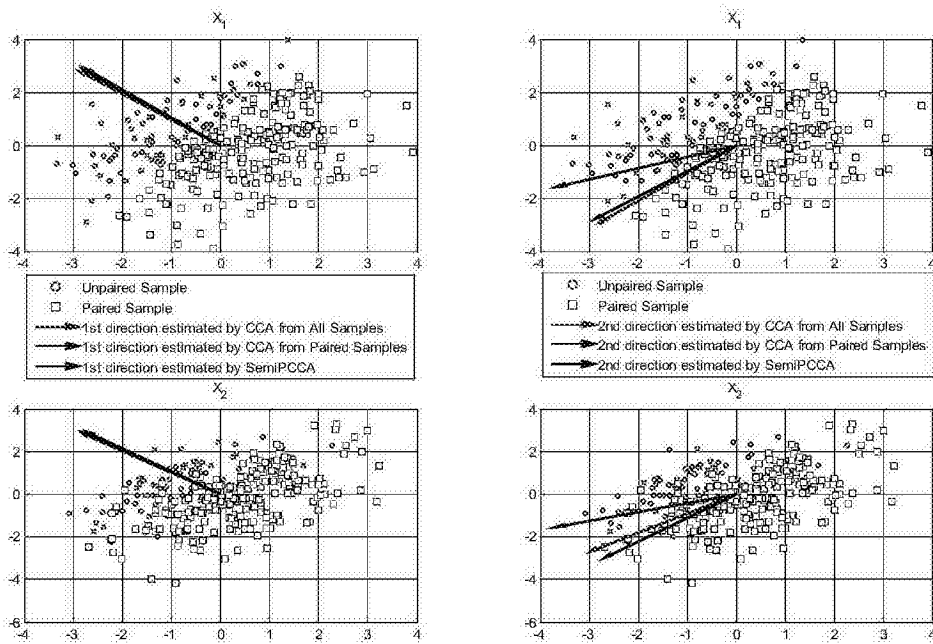


图 4a

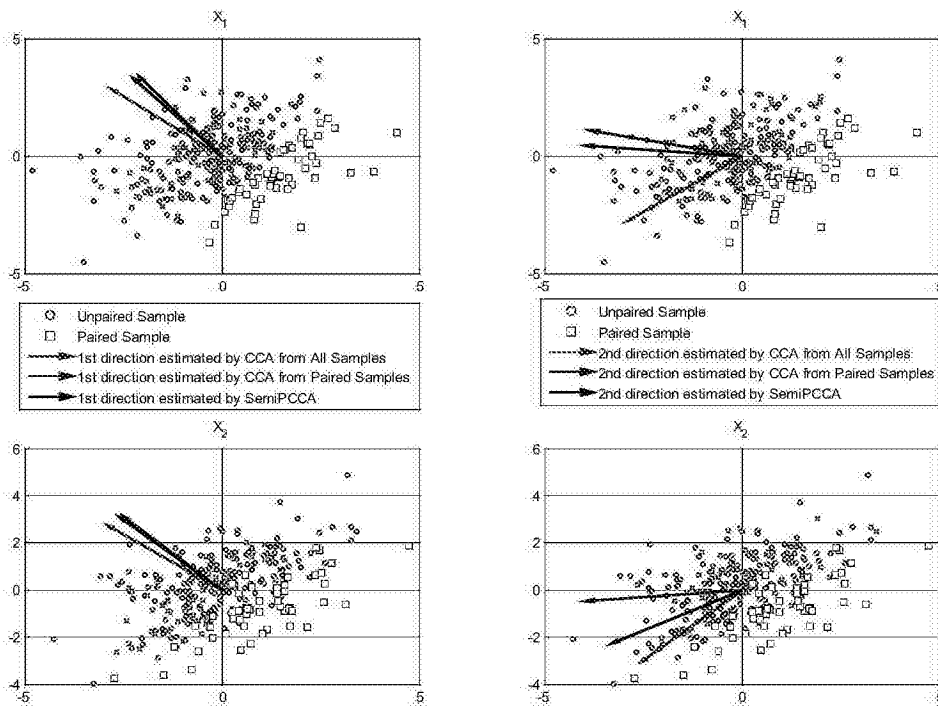
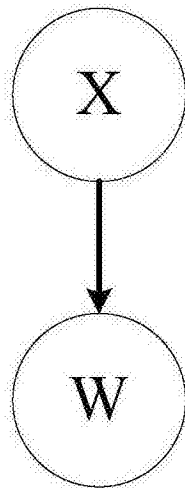


图 4b

视觉特征



语义类别

图 5a

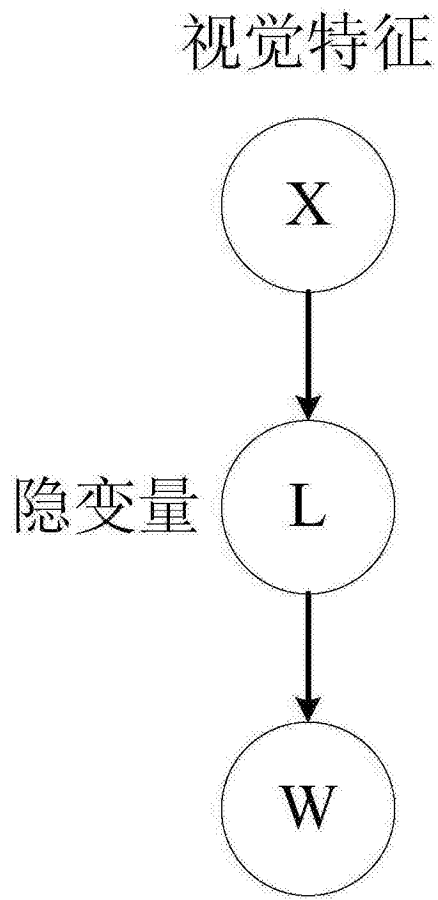


图 5b

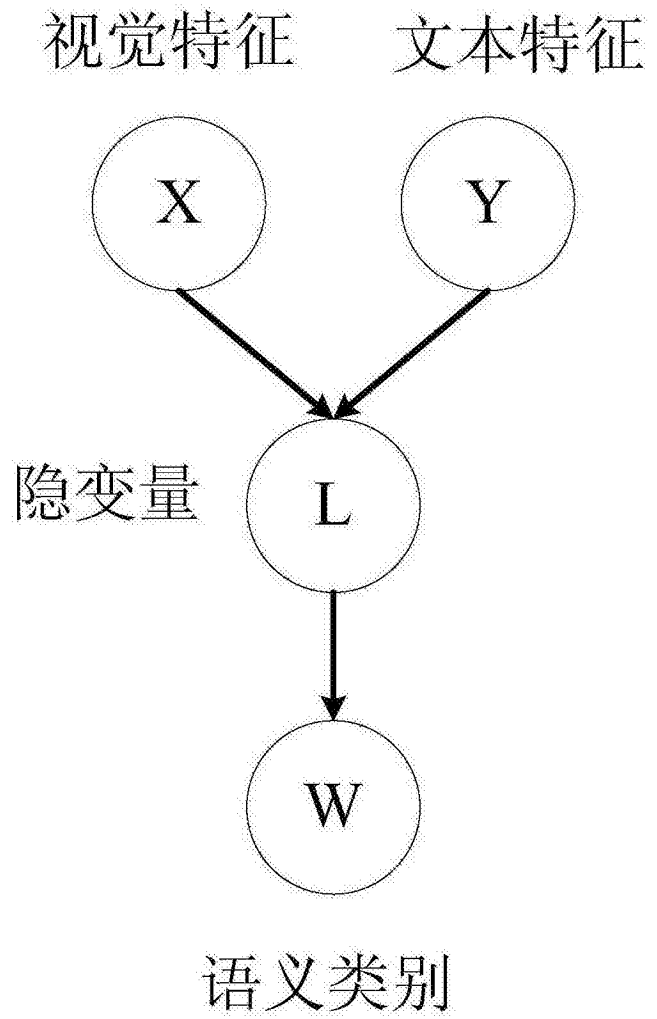


图 5c

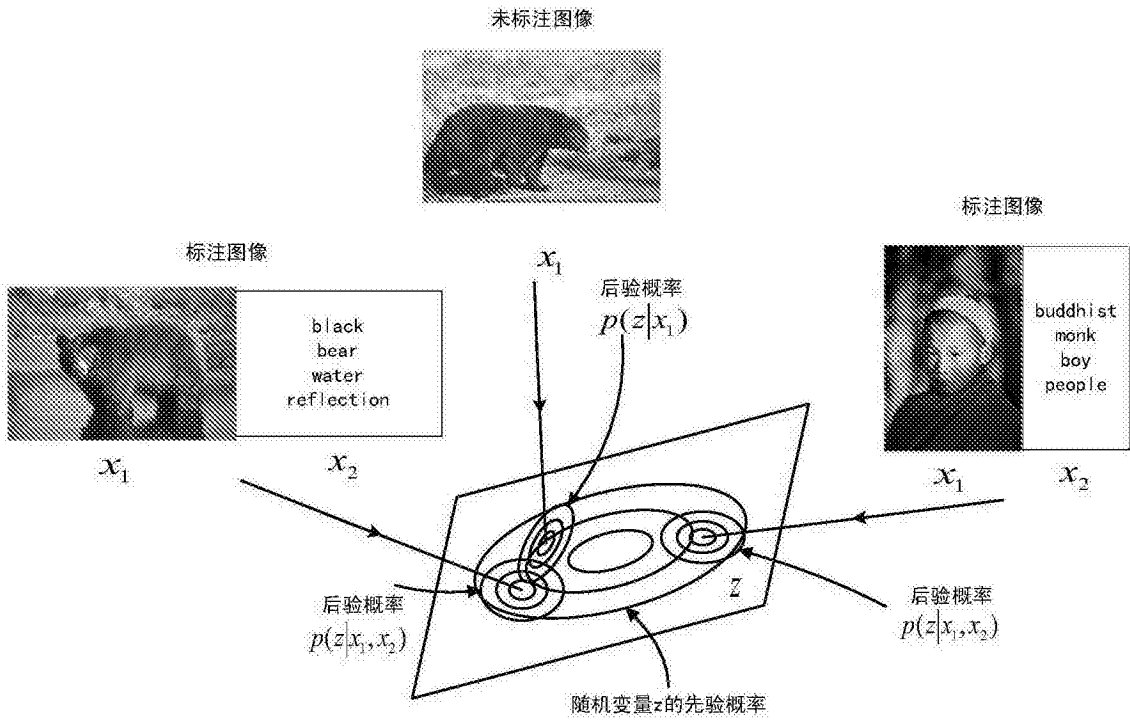


图 6

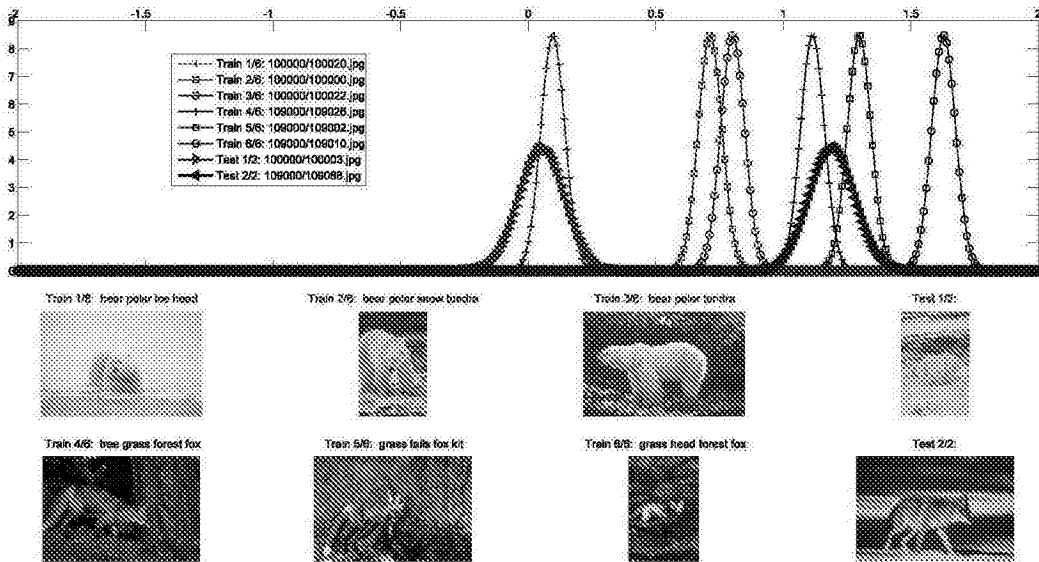


图 7

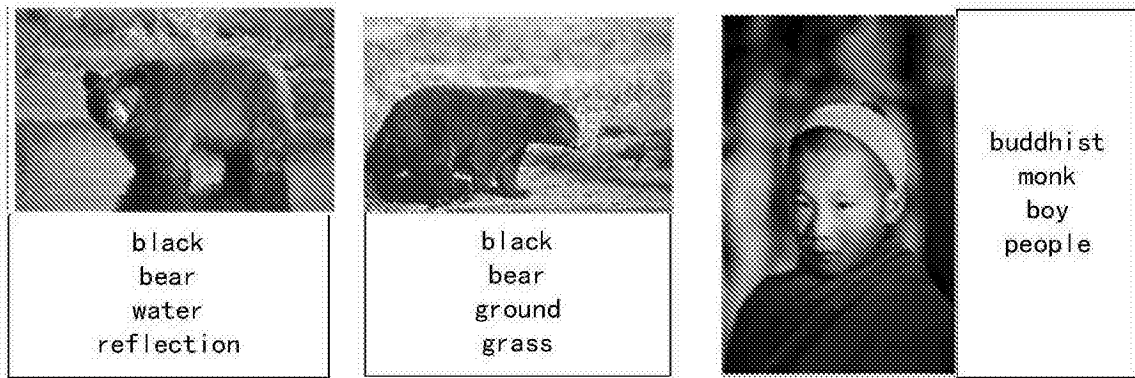


图 8

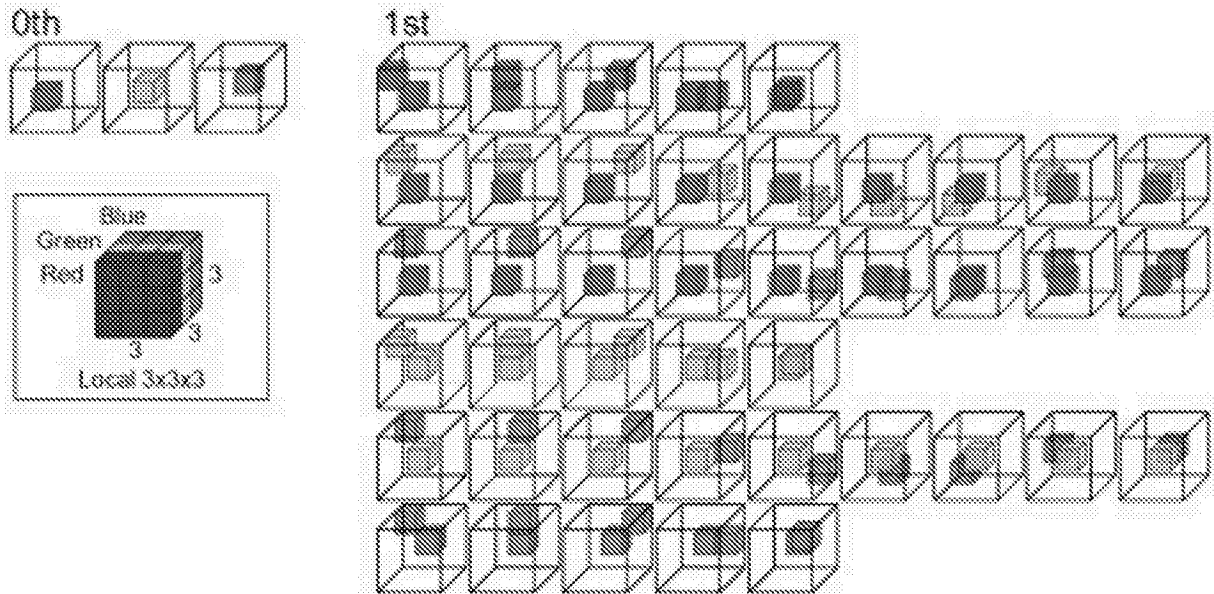


图 9