

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2017年6月8日 (08.06.2017)



(10) 国际公布号
WO 2017/092183 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2016/074646
- (22) 国际申请日: 2016年2月26日 (26.02.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510883581.2 2015年12月3日 (03.12.2015) CN
- (71) 申请人: 中山大学 (SUN YAT-SEN UNIVERSITY) [CN/CN]; 中国广东省广州市海珠区新港西路 135 号江波, Guangdong 510275 (CN)。
- (72) 发明人: 及
- (71) 申请人: 林惊 (LIN, Liang) [CN/CN]; 中国广东省广州市海珠区新港西路 135 号江波, Guangdong 510275 (CN)。 张瑞茂 (ZHANG, Ruimao) [CN/CN]; 中国广东省广州市海珠区新港西路 135 号江波, Guangdong 510275 (CN)。 王青 (WANG, Qing) [CN/CN]; 中国广东省广州市海珠区新港西路 135 号江波, Guangdong 510275 (CN)。 江波 (JIANG, Bo) [CN/CN]; 中国广东省广州市海珠区新港西路 135 号江波, Guangdong 510275 (CN)。

- (74) 代理人: 广州粤高专利商标代理有限公司 (YOGO PATENT & TRADE MARK AGENCY LIMITED COMPANY); 中国广东省广州市天河区体育西路中石化大厦 B 塔 3912 室龚素琴, Guangdong 510620 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。
- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

[见续页]

(54) Title: IMAGE RETRIEVAL METHOD BASED ON VARIABLE-LENGTH DEEP HASH LEARNING

(54) 发明名称: 一种基于可变长深度哈希学习的图像检索方法

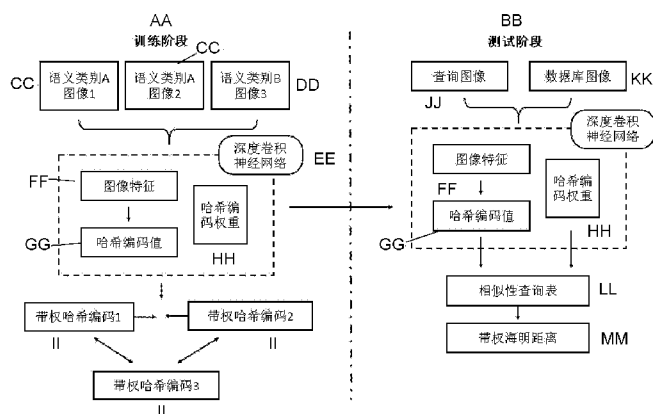


图 6

AA Training stage
BB Testing stage
CC Semantic category A image 1, 2
DD Semantic category B image 3
EE Deep convolution nerve network
FF Image feature
GG Hash code value

HH Hash code weight
II Hash code with weight
JJ Found image
KK Database image
LL Similarity searching table
MM Hamming distance with weight

(57) Abstract: An image retrieval method based on variable-length deep hash learning, which mainly relates to the fields of image retrieval and deep learning. According to the method, the learning of hash codes is modelled into a similar learning process. Specifically, according to the method, a training image is used to generate a batch of ternary image groups, wherein each of the ternary groups contains two images with the same label and one image with a different label. The purpose of model training is maximizing an interval between a matched image pair and an unmatched image pair in a Hamming space. A deep convolution nerve network is introduced into the learning part of the method, and an image feature and a hash function are optimized in a combined manner so that an end-to-end training process is realized. Furthermore, each bit of a hash code output by the convolution network has different weights. For different retrieval tasks, a user can adjust and control the length of a hash code by disconnecting an unimportant bit. At the same time, according to the method, the discriminatory property of a hash code can be effectively kept under the condition that the hash code is short.

(57) 摘要:

[见续页]

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种基于可变长度深度哈希学习的图像检索方法，主要涉及图像检索与深度学习领域。该方法将哈希码的学习建模成相似性学习的过程。具体来说，该方法利用训练图像产生一批三元图像组，其中每个三元组包含两张具有相同类标的图像与一张具有不同类标的图像。对模型训练的目的是最大化匹配的图像对与不匹配的图像对在海明空间的间隔。本方法的学习部分引入了深度卷积神经网络，通过对图像特征与哈希函数的联合优化，实现了端到端的训练过程。另一方面，卷积网络输出的哈希码每一位都具有不同的权重。针对不同的检索任务，用户可以通过截断不重要的比特位来实现哈希码长度的调控。同时，该方法能够在哈希码较短的情况下有效地保持其判别性。

一种基于可变长深度哈希学习的图像检索方法

5 技术领域

本发明涉及图像检索领域与深度学习领域，更具体的，涉及一种基于可变长深度哈希学习的图像检索方法。

背景技术

随着互联网飞速发展，以图像和视频为载体的多媒信息的数量也呈现爆炸性增长。如何才能从海量的数据中获取自己想要信息，成为工业界和学术界广泛讨论的内容。哈希技术凭借在存储空间和计算效率两个方面的巨大优势，受到了广泛的关注与研究。随着研究的不断深入，基于图像内容的哈希学习侧重将图像转化成为二进制编码后仍然能够有效保持其与其他图像的语义相关性。为了解决以上问题，多篇基于机器学习的哈希学习算法被提出。其中，一类基于有监督的图像哈希学习算法表现出了更为良好的效果，有着更加巨大的潜力。

一般来说，有监督框架下的图像哈希学习有两个主要部分构成，首先，图像将被手工设计的特征描述子表达成特征向量的形式，这一步的目的是为了在图像存在噪声或者其他冗余信息的情况下提取出图像的语义信息；其次是将哈希的学习过程表达成一个图像之间距离的优化问题，以此在学习到的海明空间中保持图像的语义一致性。但是大多数的学习方法将以上两个过程拆分开进行处理，这样就使得图像的特征表达无法在特征的学习过程中进行优化。同时，通过手工设计图像的特征需要大量的领域知识，很难有效适应不同的应用任务。近几年，一种名为多层自动编码器的模型被提出，该模型可以直接从原图像中得到图像特征并进一步转化为二进制编码。但由于其复杂的优化过程使得其在实际应用中受到很大限制。

另一方面，现有大多数的哈希学习算法生成的哈希编码是固定长度的。但是在实际的应用场景中，却需要根据任务的不同选取不同的编码长度。例如，对于计算资源较少的设备，较短的哈希编码具有更为明显的计算优势。但是对于查准率较高的任务，选择较长的哈希编码的可能性相对较大。为了解决以上问题，一个较为直接的方案是存储不同比特长度的哈希编码，但是这样对学习时间和存储

空间都有较大的浪费。因此有若干长度可变的哈希学习算法被先后提出。该类方法根据重要性的大小一位一位的产生哈希码，重要性较大的比特位率先产生。因此用户可以根据需要，从头至尾选取所期望的哈希编码长度。但是这类方法通常都与精心设计的特征空间紧密相连，同时其检索准确性也会随着哈希码位数的减少出现急剧的下跌。

发明内容

为了克服现有技术的不足，本发明提出一种基于可变长深度哈希学习的图像检索方法，该方法通过最大化匹配的图像对与不匹配的图像对在海明空间的间隔，来语义相似图像的特征一致性。通过对现有深度卷积神经网络的改进，实现了端到端的模型结构，从而能够从原始的输入图像直接获得相对应的哈希编码。最后，该方法在训练的过程中对每一个哈希比特位引入了权重，进而能够通多权重对哈希码的长度进行截取，使得哈希编码具备了长度可变性。从而实现了在一次训练的情况下，模型能够针对不同应用场景选取不同长度哈希编码进行减检索的目的。

为了实现上述目的，本发明的技术方案为：

一种基于可变长深度哈希学习的图像检索方法，包括以下步骤：

S1. 预处理：将训练图像集合划分成一批三元图像组；

S2. 训练阶段，图像哈希码的生成：将三元图像组输入到深度卷积神经网络中，通过网络的变换直接输出与图像对应的哈希码；

S3. 训练阶段，深度神经网络参数的优化：计算每张图像产生的损失，并通过反向传播算法来训练深度卷积神经网络；

S4. 测试阶段，图像相似性的计算：根据用户的输入，截断不重要的哈希比特位，并且计算查询图像与数据库每张图像之间的带权重海明距离；

S5. 测试阶段，返回查询结果：根据步骤 S4 中的带权重的海明距离，从大到小对数据库图像进行排序，排序结果即为返回的相似性检索结果。

本发明通过训练图像集合产生一系列的三元组，利用图像三元组来最大化匹配的图像对与不匹配的图像对在海明空间的间隔，进而有效地挖掘了训练图像类内的一致性和类间的差异性，解决了具有相同语义信息的图片在海明空间相似性保持的问题。

优选的，所述步骤 S1 中将训练集合划分成一批三元图像组，具体方法为：

从训练数据集中随机抽取若干张图像作为中心图像。在确定中心图像的基础上，

选取一张与中心图像具有相同类别标签的图像与一张具有不同类别标签的图像，从而形成若干图像三元组。

本发明的特点在于学习部分引入了深度卷积神经网络，通过对图像特征与哈希函数的联合优化，实现了端到端的训练过程。从而可以利用卷积神经网络直接从原始图像提取哈希码。省去了传统方法中手动设计特征的环节。其中深度神经网络的训练过程则是通过累加每张图像所产生的损失作为总的损失，进而利用反向传播算法对网络中的参数进行更新。

优选的，步骤 S2 中的深度卷积神经网络包含以下组成部分：若干卷积层和池化层，部分全连接层，一个类双曲正切函数层以及一个哈希权重层。其中倒数第二个全连接层的输出表示了图片的特征向量。倒数第一个全连接层中的参数代表了哈希函数的参数。类双曲正切函数层是一个以特征向量为输入的，每个维度的取值范围在 $[-1,1]$ 之间的类双曲正切函数，其中函数有一个调和参数用于控制其平滑性。该参数越小，函数越平滑。哈希权重层是一个以类双曲正切函数层输出为输入的，每个维度对应一个权重的深度网络层。

优选的，步骤 S3 中的深度卷积神经网络参数学习的具体实现方式为：以图像为中心来计算所产生损失，计算某图像在所有三元组中出现的损失，并将其累计求和，获得最终的图像损失。具体的，对于确定的某张图像，首先确定其在三元组中的具体存在形式（作为中心图像、作为正样本、作为负样本或是根本不在三元组中）。随后根据具体的存在形式计算其相应的损失，若图像不存在于某三元组中，则计图像在该三元组中产生的损失为零。累加图像在所有三元组中不同存在形式下的损失，获得最终该图像的损失。每次在进行深度卷积神经网络参数更新的时候则是将一批图像一次载入，利用这些图像随机产生若干三元组。统计这一批图像在这个三元组集合上的损失。利用反向传播算法（back propagation algorithm）对深度卷积神经网络参数进行更新。

本发明的优势在于针对不同查询任务，用户可以根据实际需要对哈希码的长度进行选择，

进而算法根据每个哈希比特位的权重对哈希码进行截断，实现了哈希编码的长度可变性。本发明有效地解决了一次训练，多任务可用的目标。在计算资源、存储空间有限的情况下，本方法依旧能够在哈希编码较短的情况下，保持其判别性。

优选的，步骤 S4 中的图像相似性计算的具体实现方式为：获取确定长度的哈希编码，根据用户的输入，确定要选取的哈希比特位长度 L 。利用深度卷积神经网络中哈希权重层学习到的每一个比特位的权重，从大到小依次选取权重较大的前 L 个比特位，从而实现哈希比特位的截取。利用截取的哈希比特位计算带权重的海明距离，作为最终两个图像的相似性度量。

优选的，步骤 S4 中计算带权重的海明距离使用了快速查表法，其具体的实现方式为：根据选取的哈希比特位长度 L ，构建大小为 2^L 的查询表，该查询表中的每一个元素对应了两个长度为 L 的哈希编码可能存在的一种异或操作的结果（共存在 2^L 种可能）。结合每一个哈希比特位的权重，计算带权情况下每一种异或操作的值，并将其存储在查询表的对应位置。查询时，根据输入的两个哈希编码值，计算其异或操作的结果，并将存储在查询表中对应位置的数值作为结果返回，从而获得两个哈希编码带权重的海明距离。

优选的，步骤 S5 中对带权重的海明距离进行排序，其所采用的排序算法是快速排序算法、归并排序算法或选择排序算法中的任意一种。

本发明提取具有丰富信息的图像特征和学习近似哈希函数是图像检索领域两个十分重要的研究方向。同时，当前的多媒体应用要求输出的哈希码能够依据应用场景的不同而变化长度。基于以上需求，提出一种基于可变长度深度哈希学习的图像检索方法，主要涉及图像检索与深度学习领域。该方法将哈希码的学习建模成相似性学习的过程。具体来说，该方法利用训练图像产生一批三元图像组，其中每个三元组包含两张具有相同类标的图像与一张具有不同类标的图像。对模型训练的目的在于最大化匹配的图像对与不匹配的图像对在海明空间的间隔。本方法模型的学习部分引入了深度卷积神经网络，通过对图像特征与哈希函数的联合优化，实现了端到端的训练过程。另一方面，卷积网络输出的哈希码每一位都具有不同的权重。针对不同的检索任务，用户可以通过截断不重要的比特位来实现哈希码长度的调控。同时，该方法能够在哈希码较短的情况下的有效地保持其判别性。

与现有技术相比，本发明的有益效果为：

1. 本发明是一种基于可变长深度哈希学习的图像检索方法，通过利用深度神经网络，该方法将图像的特征学习与哈希函数的学习过程统一到一个端到端的框架中，实现了两者的联合优化。克服了已有方法在哈希函数学习方面，严重依赖

于所嵌入的图像特征空间的现状。

2.本发明是一种基于可变长深度哈希学习的图像检索方法，在进行检索的过程中，该方法强调的是通过对带权重的比特位的选取来应对不同的检索任务。因而克服了大多数现有方法通过存储不同比特位的哈希码来解决不同的任务的弊端。同时，一种针对带权重哈希码的查表方法被提出，进一步保证了可变长哈希的检索效率。

3.本发明是一种基于可变长深度哈希学习的图像检索方法，在哈希码学习的过程中本方法引入了基于图像三元组比对的相对相似性比对，这种方法能够有效挖掘类内的一致性和类间的差异性。保证了所产生的哈希码能够有效保持实例级别的相似性。

附图说明

图 1 为基于可变长深度哈希学习的图像检索方法的框架图。

图 2 为本发明的基于可变长深度哈希学习的图像检索方法的深度神经网络的架构图。

图 3 为本发明的图像相似性哈希学习的原理图。

图 4 为本发明的基于可变长深度哈希学习的图像检索方法的训练流程图。

图 5 为本发明的基于可变长深度哈希学习的图像检索方法的检索流程图。

图 6 为本发明的基于可变长深度哈希学习的图像检索方法的流程图。

具体实施方式

附图仅用于示例性说明，不能理解为对本专利的限制；为了更好说明本实施例，附图某些部件会有省略、放大或缩小，并不代表实际产品的尺寸；

对于本领域技术人员来说，附图中某些公知结构及其说明可能省略是可以理解的。下面结合附图和实施例对本发明的技术方案做进一步的说明。

如图 2 所示，一种基于可变长度深度哈希学习的图像检索方法，其包括以下步骤：

预处理：将训练图像集合划分成一批三元图像组；

训练阶段图像哈希码的生成：将三元图像组输入到深度卷积神经网络中，通过网络的变换直接输出与图像对应的哈希码；

训练阶段深度神经网络参数的优化：计算每张图像产生的损失，并通过反向传播算法来训练深度卷积神经网络；

测试阶段图像相似性的计算：根据用户的输入，截断不重要的哈希比特位，并且计算查询图像与数据库每张图像之间的带权重海明距离；

测试阶段返回查询结果：根据 S4 中的带权重的海明距离，从大到小对数据库图像进行排序，排序结果即为返回的相似性检索结果。

5 上述步骤具体为：

步骤 A：将训练图像集合划分成一批三元图像组。假设训练图像可以划分为 K 个语义类，其中每个与一类包含了 P 张图像。因此总共可以产生 $K * P * (P - 1) * (K - 1) * P$ 个三元组，这个数量远远超过训练的图片数量。实际的情况下，通常训练数据集中的图片可以达到百万量级，因此在确定的硬件条件下（显存一次可以载入图像数是确定的）无法一次载入。在本实施例中，将使用分批处理的方式对深度神经网络的参数进行训练，具体实现如下：首先选择 K' 个语义类，从每一个语义类中在随机的选出 P' 张图像。对于每一张被选取的图像 I_k ，算法会生更固定个数的三元组。在每个三元组中同 I_k 有不同标签的图片是从其他语义类中进行随机选择的。利用这样的方式，在载入一批图像后，生成的三元组的语义类相对比较集中，因此训练的过程中也能够更为有效的学习到两个语义类的判别信息。但三元组的生成方式并不局限于此。

步骤 B：本实施方案能够最终实现将图像作为输入，通过深度神经网络最终输出与图像对应的哈希编码。但在实施的过程中，算法是在优化深度神经网络参数的过程中逐渐获得二值化的图像哈希编码的。具体的，本实施方案在深度神经网络中加入一层类双曲正切函数层，该层以特征向量为输入的，输出的每个维度的取值范围在 $[-1, 1]$ 之间。其形式为：

$$o(v) = \frac{1 - \exp(-\beta v)}{1 + \exp(-\beta v)}$$

其中 v 是深度神经网络倒数第一个全卷积层输出的特征表达。 β 为一个调和参数用于控制其平滑性，当 β 变得无限大的时候，类双曲正切函数将无线接近符号函数。在本实施方案中， β 的值随着训练的过程从 2 增长到 1000，因此其输出的值逐渐实现了二值化。在测试阶段，类双曲正切函数将被替换为符号函数，直接获得二值化的输出。

步骤 C：虽然本实施方案中每次产生的损失都是基于三元组计算的，但是依旧可以通过计算每张图像损失产生的梯度来进行网络参数的优化。本实施方案中

提出一种面向图像的梯度计算方法，用以优化深度神经网络的参数，但网络参数的优化方案并不局限于此。

具体实施如下：首先明确一张需要计算损失的图像 I_j ，初始化该图像的梯度为 $pd = 0$ ，对于一批载入的任意一个三元组 (I_i, I_i^+, I_i^-) ，

5 如果 $I_j = I_i$ ，则 $pd = pd + 2 * (h_i^- w - h_i^+ w)$

 如果 $I_j = I_i^+$ ，则 $pd = pd - 2 * (h_i w - h_i^+ w)$

 如果 $I_j = I_i^-$ ，则 $pd = pd + 2 * (h_i w - h_i^- w)$

 其中 h_i, h_i^+, h_i^- 是图像 I_i, I_i^+, I_i^- 对应的哈希编码， w 对应深度卷积神经网络中哈希权重层的权重。则 pd 最终的输出就为图像 I_j 的梯度。

10 步骤 D：在测试阶段，为了快速计算两张图像所产生的哈希码之间带权重的海明距离，本实施方案提出一种快速的查表法。但是，带权重的海明距离的求解形式不局限于此。

 快速的查表法具体如下：获取确定长度的哈希编码，根据用户的输入，确定要选取的哈希比特位长度 L 。利用深度卷积神经网络中哈希权重层学习到的每一个比特位的权重，从大到小依次选取权重较大的前 L 个比特位，从而实现哈希比特位的截取。根据选取的哈希比特位长度 L ，构建大小为 2^L 的查询表，该查询表的中的每一个元素对应了两个长度为 L 的哈希编码可能存在的一种异或操作的结果（共存在 2^L 种可能）。结合每一个哈希比特位的权重，计算带权情况下每一种异或操作的值，并将其存储在查询表的对应位置。查询时，根据输入的两个哈希编码值，计算其异或操作的结果，并将存储在查询表中对应位置的数值作为结果返回，从而获得两个哈希编码带权重的海明距离。在上述情况下，如果选取的哈希编码长度过长，查询表的长度将变得过大，利用查询变穷举所有可能的异或操作变得不现实。鉴于此，本实施方案提出一种分段计算方案。具体的，算法将选取的 L 个比特位分成若干段（例如 8 比特位为一段）。每一段哈希编码对应一个子表，子表中的每个位置对应了两个哈希编码对应部分异或操作的输出值。因此，总的带权重的海明距离可以通过累加每一段的返回值最终获得。

15

20

25

 步骤 E：测试的最后阶段是将数据库中每张图像与查询图像之间的带权海明距离进行大到小的排序，最终获得排序列表。在排序列表中靠前的说明和查询图像相似，靠后则说明不相似。本实施例中，使用了快速排序对带权的海明距离进行排序，但排序算法并不局限于此算法。

30

显然，本发明的上述实施例仅仅是为清楚地说明本发明所作的举例，而并非是对本发明的实施方式的限定。对于所属领域的普通技术人员来说，在上述说明的基础上还可以做出其它不同形式的变化或变动。这里无需也无法对所有的实施方式予以穷举。凡在本发明的精神和原则之内所作的任何修改、等同替换和改进

5 等，均应包含在本发明权利要求的保护范围之内。

权利要求书

1、一种基于可变长深度哈希学习的图像检索方法，其特征在于，包括：

S1. 将训练图像集合划分成一批三元图像组；

5 S2. 将三元图像组输入到深度卷积神经网络中；

S3. 计算每张图像产生的损失，并通过反向传播算法来训练深度卷积神经网络；

S4. 根据用户的输入，截断不重要的哈希比特位，并且计算查询图像与数据库每张图像之间的带权重海明距离；

10 S5. 根据步骤 S4 中的带权重的海明距离，从小到大对数据库图像进行排序，排序结果即为返回的相似性检索结果。

2、根据权利要求 1 所述的基于可变长深度哈希学习的图像检索方法，其特征在于，所述步骤 S1 中所述的三元图像组具体包括两张具有相同类别标签的图像与一张具有不同类别标签的图像。

15 3、根据权利要求 1 所述的基于可变长深度哈希学习的图像检索方法，其特征在于，所述步骤 S2 中的深度卷积神经网络包含若干卷积层和池化层，部分全连接层，一个类双曲正切函数层以及一个哈希权重层；

所述的类双曲正切函数层是一个以特征向量为输入的，每个维度的取值范围在 $[-1,1]$ 之间的类双曲正切函数，其中函数有一个调和参数用于控制其平滑性；

20 该参数越小，函数越平滑；

所述的哈希权重层是以类双曲正切函数层的输出作为输入，每个维度对应一个权重的深度网络层。

4、根据权利要求 3 所述的基于可变长深度哈希学习的图像检索方法，其特征在于，所述步骤 S3 中的深度卷积神经网络的训练过程是一个端到端的过程，

25 实现了图像特征与哈希函数的联合优化；

所述的图像特征是深度卷积网络倒数第二个全连接层输入的特征向量；

所述的哈希函数，其参数直接对应了最后一个全连接层所包含的所有参数。

5、根据权利要求 1 所述的基于可变长深度哈希学习的图像检索方法，其特征在于，所述步骤 S3 中计算每张图像的损失，是通过最大化匹配的图像对与不匹

30 配的图像对在海明空间的间隔实现的，满足：

$$W^* = \min_W \sum_i \max\{H(h_i, h_i^+) - H(h_i, h_i^-), C\}$$

其中 W 是深度卷积神经网络的参数, $H(\cdot, \cdot)$ 表示两个哈希码在海明空间的距离, C 是一个常数用于减小噪声对模型的影响。

6、根据权利要求 1 所述的基于可变长深度哈希学习的图像检索方法, 其特征在于, 所述步骤 S4 中不重要的哈希比特是通过哈希比特的权重确定的, 其中权重越小表明哈希比特位的重要性越小。

7、根据权利要求 1 所述的基于可变长深度哈希学习的图像检索方法, 其特征在于, 所述步骤 S4 中计算带权重的海明距离, 包含以下步骤:

S-a. 获取重要的哈希比特位的权重;

10 S-b. 构建一个长度为 2^l 的查询表, 其中 l 是重要哈希比特位的长度, 数值 2^l 等于两个哈希码所产生的所有可能的异或结果;

S-c. 计算每种异或结果下的带权海明仿射距离并将结果储存在查询表的对应位置上。

S-d. 计算两个哈希码的异或值并将其在查询表中对应的值返回。

15 8、根据权利要求 7 所述的计算带权重的海明距离, 其特征在于, 所述长度为 2^l 的查询表在 l 较大的情况下, 可以将查询表拆分成等长的几个子表, 每一个子表对应了固定长度的一段哈希码。子表中的每个值代表了对应子哈希码的带权相似性。最终两个哈希码的带权仿射距离可以通过累加每段的带权仿射获得。

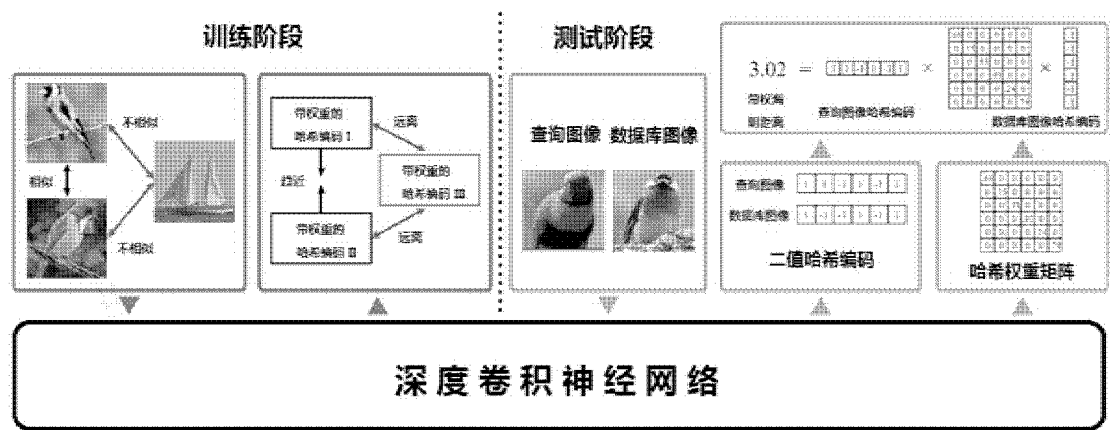


图 1

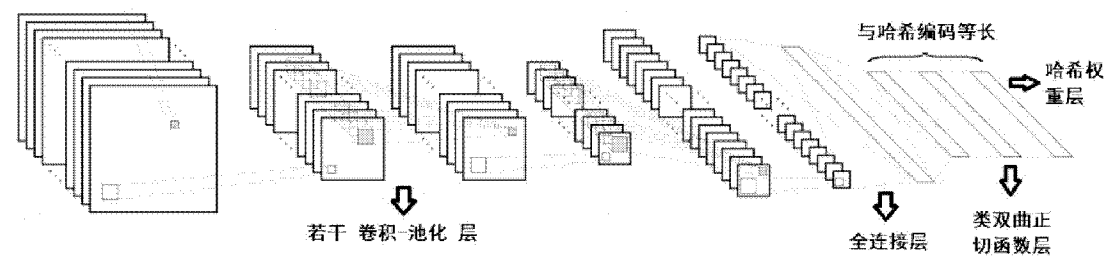


图 2

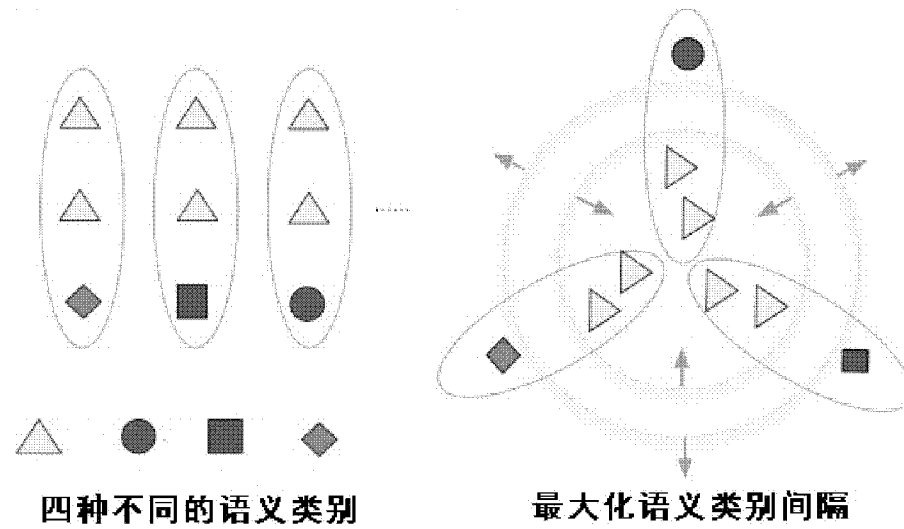


图 3

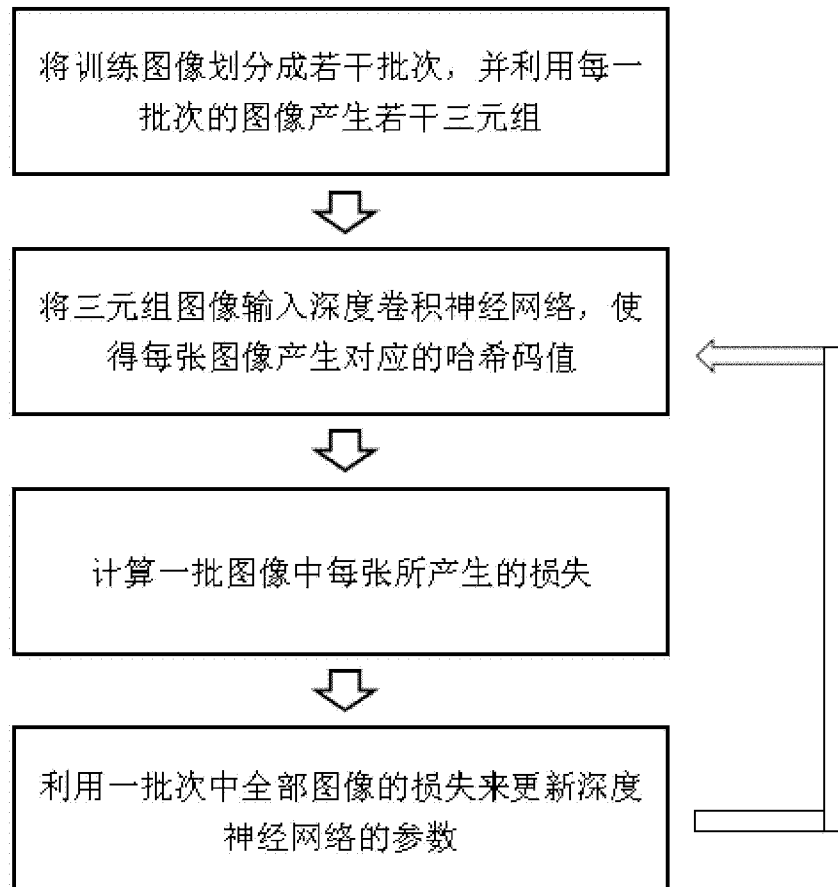


图 4

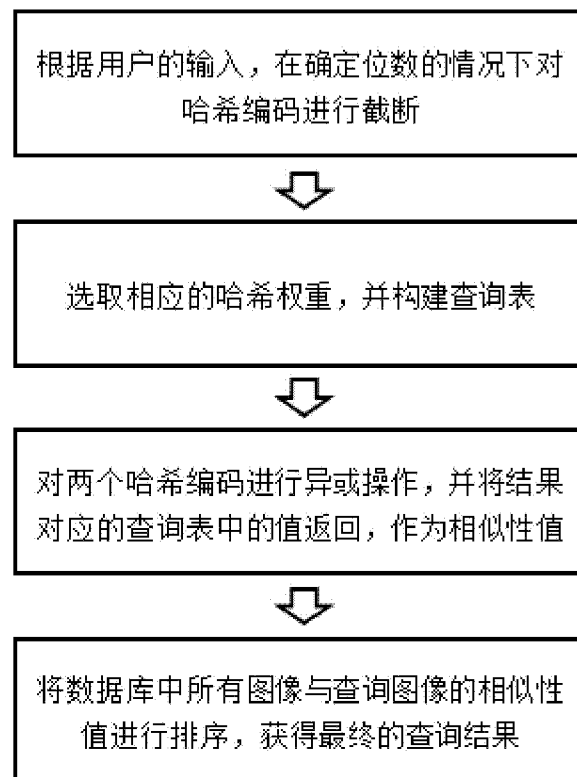


图 5

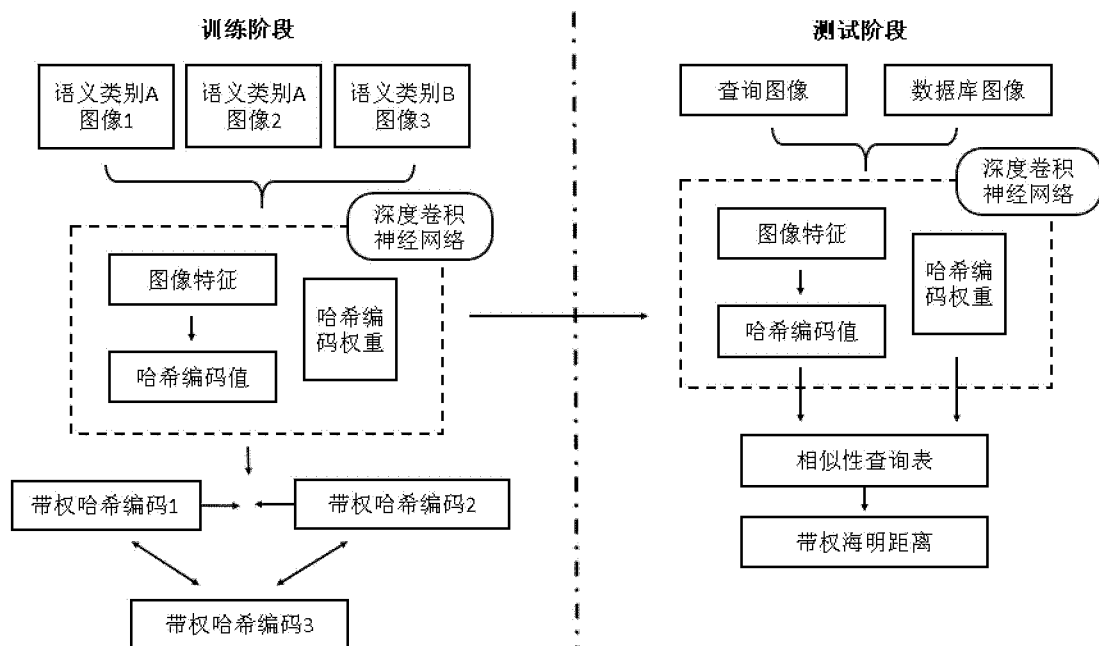


图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/074646

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI, IEEE, GOOGLE: image?, picture?, search+, retriev+, hamming, distance, hash, deep+, learn+, scale, variable, triplet?, bit, drill, convolution, nerve w network, NN, layer?, weight+, cut+, sort

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	ZHANG, Ruimao et al., "Bit-Scalable Deep Hashing With Regularized Similarity Learning for Image Retrieval and Person Re-Identification" IEEE TRANSACTIONS ON IMAGE PROCESSING, volume 24, number 12, 31 December 2015 (31.12.2015), page 4767, the left column, paragraph 1 to page 4771, the right coulmn, paragraph 1	1-8
A	CN 104834748 A (INSTITUTE OF AUTOMATION, CHINESE ACADEMY OF SCIENCE) 12 August 2015 (12.08.2015) claims 1-7, and description, paragraphs [0020]-[0054], and figures 1-3	1-8
A	US 8352494 B1 (GOOGLE INC.) 08 January 2013 (08.01.2013) the whole document	1-8
A	CN 105095435 A (BEIJING JINGDONG SHANGKE INFORMATION TECHNOLOGY CO., LTD. et al.) 25 November 2015 (25.11.2015) the whole document	1-8
A	CN 105069173 A (TIANJIN ZHONGKE INTELLIGENT IDENTIFICATION INDUSTRY TECHNOLOGY RESEARCH INSTITUTE CO., LTD.) 18 November 2015 (18.11.2015) the whole document	1-8

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 17 August 2016	Date of mailing of the international search report 26 August 2016
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer LI, Min Telephone No. (86-10) 62413700

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/CN2016/074646

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104834748 A	12 August 2015	None	
US 8352494 B1	08 January 2013	US 2015161178 A1	11 June 2015
CN 105095435 A	25 November 2015	None	
CN 105069173 A	18 November 2015	None	

A. 主题的分类 G06F 17/30 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类		
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) WPI, EPODOC, CNPAT, CNKI, IEEE, GOOGLE, 万方: 图像, 检索, 图片, 相片, 搜索, 查询, 深度, 哈希, 学习, 变长, 三元组, 卷积, 神经网络, 反向传播, 比特, 海明距离, 汉明距离, 相似性, 标签, 层, 权重, 截断, 优化, 分类, image?, picture?, search+, retriev+, hamming, distance, hash, deep+, learn+, scale, variable, triplet?, bit, drill, convolution, nerve w network, NN, layer?, weight+, cut+, sort		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	ZHANG, Ruimao 等. "Bit-Scalable Deep Hashing With Regularized Similarity Learning for Image Retrieval and Person Re-Identification" IEEE TRANSACTIONS ON IMAGE PROCESSING, 第24卷, 第12期, 2015年 12月 31日 (2015 - 12 - 31), 第4767页左栏第1段-第4771页右栏第1段	1-8
A	CN 104834748 A (中国科学院自动化研究所) 2015年 8月 12日 (2015 - 08 - 12) 权利要求1-7, 说明书第[0020]-[0054]段, 附图1-3	1-8
A	US 8352494 B1 (GOOGLE INC.) 2013年 1月 8日 (2013 - 01 - 08) 全文	1-8
A	CN 105095435 A (北京京东尚科信息技术有限公司 等) 2015年 11月 25日 (2015 - 11 - 25) 全文	1-8
A	CN 105069173 A (天津中科智能识别产业技术研究院有限公司) 2015年 11月 18日 (2015 - 11 - 18) 全文	1-8
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期 2016年 8月 17日		国际检索报告邮寄日期 2016年 8月 26日
ISA/CN的名称和邮寄地址 中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 李敏 电话号码 (86-10)62413700

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/074646

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	104834748	A	2015年 8月 12日	无	
US	8352494	B1	2013年 1月 8日	US 2015161178 A1	2015年 6月 11日
CN	105095435	A	2015年 11月 25日	无	
CN	105069173	A	2015年 11月 18日	无	