基于云数据的高效图像编码方法

赵琛[1)]，马思伟[1)]，张新峰[2)]，张健[1)]，高文[1)]

[1)](北京大学 信息科学技术学院，北京，100871)

[2)](南洋理工大学 博云搜索， 新加坡，637553)

摘 要 随着数字图像采集设备的普及，加之互联网技术的高速发展，海量的图像在上传并保存到互联网中。传统的图像编码方法（如 JPEG, JPEG2000），可以在一定程度上压缩图像数据，但其独立压缩单幅图像的方式导致其压缩效率远远不能满足当今大数据的需求。为了更高效地压缩互联网图像数据，我们充分利用“大数据”这一挑战，提出一种基于云数据的图像编码方法。针对上传到云数据库中的图像，算法采用“上传一幅、压缩一幅”的处理模式，有效地适应迅速变化的大数据环境。它充分利用了当前要图像和云数据中其它图像之间的相关性，更加有效地对图像进行压缩编码。在该方法中，我们有三个创新性技术设计。首先，基于高效的图像检索技术，设计了一个方案对“大数据”进行筛选，仅保留与当前图像高度匹配的少部分图像数据。其次，利用这些经过选择得到的少量图像数据，我们进行图像间的预测编码。这一环节对于整个系统的性能非常重要。我们采用基于块匹配的预测编码和率失真优化技术，在保证图像重建质量的前提下将码率大大减少。还有，我们设计了一套对检索得到图像的预处理方案，将其变换成更为接近于当前图像的参考图像，从而提高预测准确度。我们采用投影变换技术将检所得到的图像与当前图像对齐，并通过光照补偿使得两幅图像中的匹配像素点的值尽可能接近。本文为大数据下的图像压缩提供了一个崭新的框架，可以大大节省如网络云盘中图像所占用的存储空间。相比于传统图像编码方法甚至是更为高效的 HEVC 帧内编码方案，该图像压缩系统都能取得明显的性能提升。实验表明，该方法对图像的压缩效率，比 JPEG 平均提升 78.5%，比 HEVC 帧内编码能提升 67.2%。甚至相比于文献中最新的基于云的图像编码方法，本文的算法在同码率下的主观质量和客观质量都表现出明显优势。

关键词 大数据；云编码；图像压缩；HEVC；图像检索；投影变换；光照补偿

中图法分类号 TP391

An efficient image compression method based on cloud data

ZHAO Chen[1)], MA Si-Wei[1)], ZHANG Xin-Feng[2)], ZHANG Jian[1)], GAO Wen[1)]

[1)](School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China)

[2)](Rapid-Rich Object Search Lab, Nanyang Technological University, Singapore 637553)

Abstract With the rapid development of the Internet technology and the digital imaging device, massive pictures are uploaded and saved in the network disks every day. Traditional image coding methods (such as JPEG, JPEG2000) only exploit spatial correlations inside a picture, thus their coding efficiency cannot satisfy the demand of the huge picture data. In order to compress the images in the Internet more efficiently, we proposed a novel cloud-based image compression method by taking advantage of the ‘big data’ challenge, rather than simply treating it as a difficulty. The proposed system adopts the processing mode of ‘encoding one image immediately after it is uploaded’, which effectively adapts to the rapid change of the big-data environment. It compresses an

本课题得到国家重点基础研究发展计划(973)项目 (No. 2015CB351800)、国家自然科学基金优秀青年科学基金(No.61322106)和国家自然科学基金（No. 61370114)的资助. 赵琛(通讯作者)，女，1988年生，博士研究生，主要研究领域为图像视频处理、图像视频编码，E-mail: zhaochen@pku.edu.cn. 马思伟，男，1979年生，博士，教授，主要研究领域为图像视频编码、视频处理、视频传输，E-mail: swma@pku.edu.cn. 张新峰，男，1983年生，博士，博士后，主要研究领域为图像视频处理、图像视频编码. 张健，1985年生，博士，博士后，主要研究领域为多媒体图像视频压缩及处理. 高文，1956年生，博士，教授，中国科学院院士，主要研究领域为图像处理、视频编码和传输、模式识别、多媒体信息检索、多模式接口等.

image by fully exploiting the correlations between the image-to-be-compressed and other images in the cloud. We had three innovative technical designs to serve for the compression framework. a) We designed a scheme based on an efficient image retrieval technique to effectively filter the 'big data' and to only keep the 'small data' that well match the image to be compressed. b) By utilizing the selected 'small data', we did inter-image prediction coding, which is essential for the performance of the compression system. We exploited the techniques of block matching-based prediction coding and rate-distortion optimization, so that coding bits are tremendously reduced in the premise of good reconstruction quality. c) We ensured the prediction accuracy by applying a couple of preprocessing techniques on the retrieved similar image to obtain good references. In this step, we applied projective transformation to align the retrieved image and the current image, and did illumination compensation to make the values of corresponding pixels as close as possible. This paper presents a novel framework for image compression in the big-data era, which tremendously saves storage requirement for images in the cloud disks. The proposed system has significant compression performance gains compared to traditional image coding methods and even the advanced HEVC intra coding. The experimental results demonstrate that it outperforms JPEG and HEVC intra coding by 78.5% and 67.2% on average, respectively. Moreover, compared to the existing state-of-the-art attempt for cloud-based image coding in the literature, both the objective and subjective reconstruction quality of our algorithm has obvious improvement.

Key words big data；cloud-based image coding；image compression；HEVC；image retrieval；photometric transformation; illumination compensation

1 引言

近年来，随着智能手机、数码相机的日益普及，大量的图片数据被采集和存储。随之而来的，是人们对存储卡、硬盘容量要求的不断升级。个人硬盘，已经从几年前的以 GB(1GB=1024MB)为单位的存储容量，上升到以 TB(1TB=1024GB)为单位。而且，依照目前的图片数据增长趋势，这将很快不能满足用户需求。为了解决个人存储空间不足的问题，由互联网企业提供的云存储服务应运而生，如百度云盘、华为网盘、人人相册、QQ 空间相册。但这只是将存储的压力从个体用户转向了企业用户，并不能从根本上解决问题。据《2014 年 3 月中国个人云存储市场用户数（网页版）》数据显示[1]，国内个人云存储用户数（网页版）已超过 3.51 亿人。 根据 2012 年腾讯公司的调查数据，一天之内，Facebook 就有 2.5 亿张图片上传[2]。数据量的飞速增长，不仅对硬件存储提出了越来越高的要求，对能源消耗、人力成本也是一项严峻的考验。

为了有效地存储海量的图像信息，需要对其进行压缩编码。传统的图像编码方法（如 JPEG，JPEG2000），利用变换、量化和熵编码等方式，可以去除单幅图片内部的信息冗余。但这些编码方法只利用到图像内部的相关性，压缩率远远不能满足当今大数据时代海量图片数据的存储需求。在当今时代下，数码拍照设备的普及产生了大量的图片，这既是硬件存储的一项挑战，同时也提供了更进一步压缩数据的可能性。由于拍照成本的降低，人们会习惯于在同一地点、对于同一物体拍摄多张照片，这些照片可能有光照、人物表情姿势、拍摄角度等的差异，但景物内容的相同使得图片之间存在非常大相似性。不仅是同一相册的照片具有冗余性，不同人在同一地点拍摄的照片同样具有相关性。旅游景区每年都会吸引大量的游客，游客们会在地标性建筑前留下多张纪念照片，这样，人们上传于云端的海量照片之间存在大量信息冗余。因此，图像之间的信息相关性可以加以利用，来进一步提升压缩效率。

1 http://info.hhczy.com/article/20140606/21626.shtml

2 http://tech.qq.com/a/20120306/000306_1.htm

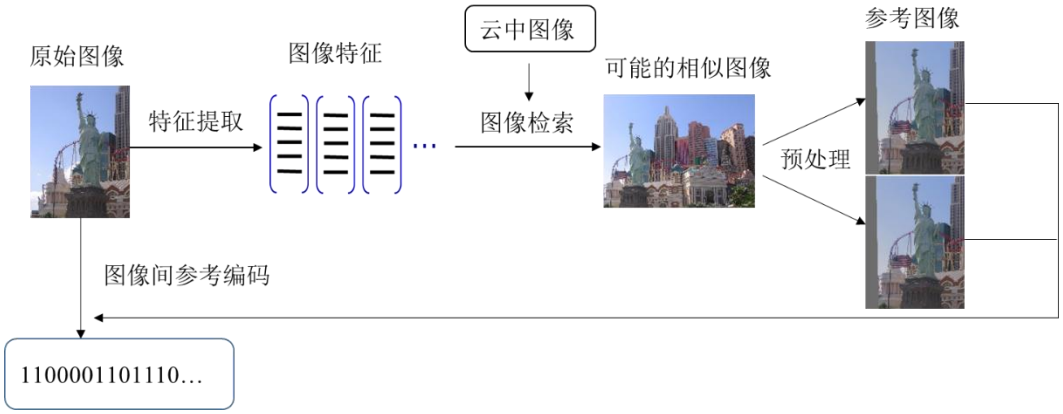


图 1 基于云数据的图像编码算法框架图

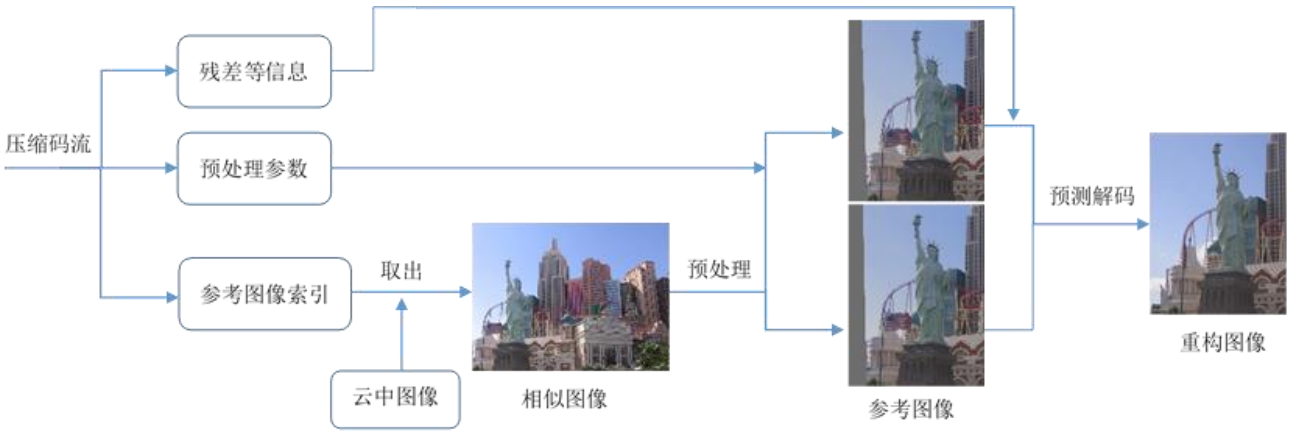


图 2 基于云数据的图像解码算法框架图

利用图像间相关性进行压缩编码，人们早在上世纪 90 年代就进行了探索，他们针对图像集或者相册的编码，提出了有效的算法[1]-[12]。这些算法，试图寻找一个合理的图像组织方式，通过图像间的预测对不同图像的冗余信息进行压缩。至今仍有学者在这方面做出不错的工作，能够更高效地针对相册中的图像进行压缩。但这些图像集压缩的方法，大多利用所有图像整体计算参考方式，并不适于直接应用于当前大数据的压缩。首先，云端具有海量的图片数据，对所有的图片整体计算出一种最优参考方式的复杂度非常高，在实际应用中并不现实；其次，云端每分每秒都会有新的图片上传，所以并不能一次性获得所有图片信息进行计算。所以，我们需要针对当前的应用环境，设计出新的图像编码方式，来更有效地利用云数据实现更高效压缩。

Huanjing Yue 等学者，首次系统地提出了一种基于云平台的图像编码方案[13]。在该方案中，编码端对原始图像进行下采样，得到一个分辨率较低的小图像；同时，提取图像特征对其进行描述，并利用小图像对这些特征信息进行压缩。小图像和特征信息经过编码后，被传送到解码端。在解码端，首先解码小图像，在它的指导下对特征信息进行解码。然后，利用这些解码得到的图像特征在云端搜索相似图像块，将它们拼接重构出原始图像。在该方法中，不需要存储和传输原始图像的全部信息，而是只需保留压缩后的特征及下采样后的小图像，从而在很大程度上减少了码流大小。但这种方案假定在云端一定能够检索到相似图像或者跟当前图像几乎一样的图像，并没有为云中不存在相似图像的情况提供补偿编码机制。一旦找不到相似图像，或者找到的图像块和原始图像块差异很大，那么重构图像的质量将非常低，甚至重构图像会出现严重的扭曲变形。而且，该方法是通过块的拼接来重建原始图像，码流中没有保留任何残差信息。所以，即使在相似图像存在的情况下，它得到的重构图像即使视觉上看似接近于原始图像，客观保真度也难以保证，甚至低于传统的 JPEG。

为了解决大数据环境下的图像压缩问题，本文

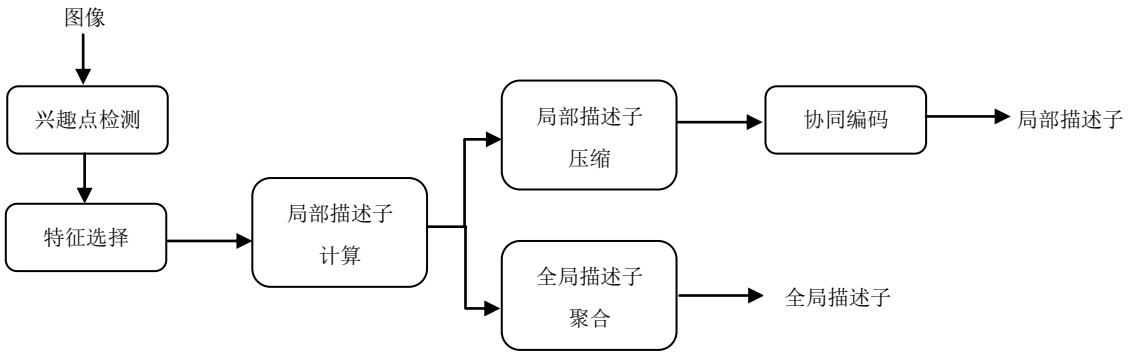


图 3 视觉搜索紧凑描述子（CDVS）生成流程图

提出了一种新型图像编码算法，它针对用户新上传到网络云盘(云数据库)中的图像进行压缩，目的在于利用云盘中已有的大量图片数据，来提高当前图像的压缩效率。与图像集编码思想不同，该算法不需要计算云中所有图像的结构，而是针对每次新上传到云端中的一幅图像进行编码，即采用“上传一幅、压缩一幅”的处理模式，保证了在大数据下的时间可行性。同时，相比[13]的编码方案，该算法在编码端而非解码端进行相似图像检索，不仅大大节省了解码时间，而且使得在未找到相似图像时可采用图像内编码方式，从而保证了图像的客观保真度。

以下的章节安排如下：第 2 部分总述本文提出的基于云数据的图像编码方法，给出编解码技术框架。第 3 部分对特征提取和相应的图像检索方案进行详细介绍。第 4 部分讨论了对相似图像的预处理方案。第 5 部分针对图像间预测编码进行探讨。第 6 部分提供作者的仿真实验结果，并给出相应的分析与讨论。第 7 部分对文章进行总结。

2 基于云数据的图像编码

我们假定在当前网络云盘(云数据库)中已有大量图像，这些图像既可以是以如 JPEG 等传统方式编码，也可以是以[1]-[12]中所述的某种图像集方式编码（要求参考层数不要太大，能保证快速解码）。由于数据库中的图片数目非常大，所以很可能会有图片与下一张要传入云盘的图片具有相似性。这样，云盘中与当前图像相似的图片，经过解码，就可以为当前图片的压缩提供参考；对当前图片的压缩，仅需要在参考图片基础上再额外附加一些数据（如残差、参考索引），省去了压缩整幅图片所用的大量比特数。

基于这一想法，本文提出一个基于云数据的图像压缩方法及相应的编解码系统。它利用云数据库中已有的图像信息，对用户上传的每一幅图像按照“上传一幅、压缩一幅”的模式进行编码，即用户每上传一幅图像，系统便只针对这一幅图像进行压缩处理，而不需对云图像库中的所有图像进行计算（这与图像集编码中获得所有图像之后再集体编码不同）。在本文中，我们将编码对象---新上传至云中的图像，称之为当前图像。

图 1 表示了本文算法对当前图像的编码过程。整个编码过程分为四个关键步骤：对当前图像的原始图像进行特征提取；利用特征在云数据库中检索得到可能的相似图像；对检索到的图像进行预处理，得到多幅参考图像；根据参考图像和当前图像之间的相关性，进行图像间的预测编码。因此，在该系统中，需要编入码流的数据仅仅是图像间预测编码之后的残差及参考信息等数据、相似图像的索引和预处理的参数，这些数据量要远小于直接编码单幅图像所消耗的比特数。

解码过程如图 2 所示。首先，根据压缩码流中包含的参考图像的索引信息，从云图像数据库中取出编码端检索到的相似图像。然后，根据码流中的预处理参数，对相似图像进行预处理，这样得到跟解码端一样的参考图像。最后，利用参考图像进行预测解码，结合残差信息和预测图像块得到重构的图像。可以看出，解码过程比编码要简单很多，它既不需要像编码端一样进行图像检索，也不需要复杂的图像间参考的块匹配过程，因此能做到实时解码。

3 特征提取及相似图像检索

在本章节中，我们介绍本文所提出的图像编码系统中图像特征的提取，以及利用这些特征在云数据库中检索相似图像的方法。

3.1 CDVS特征提取

云数据库中存有海量图片，这为当前图片相似图片的存在提供了很大的可能性；但同时并非云中的每幅图片对当前图片的编码都有帮助，所以我们需要从海量的云图片中检索到与当前图像最相近的一幅图像。当前图像的相似图像，定义为与它的拍摄对象或拍摄场景相同，即在图像内容上具有很大的一致性的图像。

相似图像的查找，在图像集编码的文献中提供了多种方法[1]-[12]，同时，也有一些专门针对图像检索所提出的图像特征及匹配方法 [14]-[16]。其中典型的方法有，从图像中提取尺度不变特征变换（Scale-invariant feature transform，SIFT）特征点，然后进行图像间的特征点匹配，通过匹配特征点距离的计算来确定相似图像[1]。每个 SIFT 特征点包含一个 128 维的描述子；平均来说，每个百万像素级的图片又有上千个 SIFT 特征点；而云图像数据库中至少有上千幅图片。这样，每次图像检索，都需要对当前图像和云图像库中的每个图像进行特征匹配，即对当前图像中的每个特征点和库中每幅图像的每个特征点进行匹配，然后根据匹配特征点的距离再计算当前图像和库中每幅图像的距离。这种方式在图像数量较少时比较有效，但在海量图片中进行查找，它需要耗费大量的运算时间，不适用于云数据下的图像编码。

考虑到计算复杂度，在我们的图像编码框架中，不采用 SIFT 特征点进行图像检索，而是利用视觉搜索紧凑描述子（the MPEG standardized compact descriptor for visual search，CDVS）[17]。CDVS，是 MPEG 组织对图像检索的描述子格式、特征提取和检索过程所制定的标准。相比于一般的图像特征，它不仅占用字节数更少（一幅图像只有几 KB 甚至更少），具有非常好的检索和匹配性能，而且特征提取和匹配非常快。

CDVS 的生成流程如图 3 所示。首先，对一幅图像进行兴趣点检测，称之为局部特征点。然后，从所有的局部特征点中进行筛选，去除掉噪声点，只保留重要的局部特征点。以这种方式，从局部特征点的数目上实现了数据的压缩。接下来，对保留的局部描述子，进行两个方面的处理：局部描述子的压缩得到更紧凑的描述子；局部描述子的聚合得到全局描述子。局部描述子经过变换和标量量化，实现了数据量的压缩。而对它进行降维、聚合以及二值化处理之后，便可形成全局描述子。可以看出，从图像中检测选择得到的特征点，最后形成了两部分的特征信息：局部描述子和全局描述子。对局部描述子的压缩，使得 CDVS 特征所占用的字节数相比原始兴趣点大大减少，从而特征匹配所用的时间也相应减少。同时，全局描述子的生成，使得图像有不同层次的信息描述，从而提高了检索的准确性。

3.2 基于CDVS特征的相似图像检索

利用 CDVS 特征进行图像检索，也同样基于全局描述子和局部描述子，如图 4 所示。首先，对数据库中的所有图像，基于多块索引结构生成一个全局描述子的索引表，由此产生有待匹配的候选图像。这种方式大大减少了后续特征匹配所需要的计算量，提高了检索效率。将当前图像和候选图像的全局描述子，基于汉明距离进行匹配，根据匹配结果选出排名最靠前的若干个（如 500 个）图像，这样进一步缩小了候选图像范围。最后，利用局部描述子对所有候选图像进行重排序，排序越靠前的图像表示与当前图像的相似度越高。

根据检索图像的排列顺序，我们从中选取排在第一位的图像作为可能的相似图像。这个相似图像在整个云图像数据库中的索引号，经过编码并传输到解码端。如果这个检索到的相似图像确实是跟当前图像具有相同拍摄内容，那么经过我们下一步的图像预处理，它将与当前图像对应像素点的值更为接近，从而为当前图像提供非常好的预测。如果这个图像是一个错误的检索结果（可能由于云中根本没有当前图像的相似图像，也可能云中存在但没有检索到），我们的系统在最后预测编码阶段将自适应地选择编码方式，切换为图像内的预测方式进行编码。

4 相似图像预处理

本章节将介绍本文图像编码系统的第三个环节，参考图像的预处理。这一过程，对上一环节中检索得到的相似图像进行处理，使其在物体位置、图像亮度等方面与当前图像更为接近，从而为进一步的以图像块为单位的预测做好准备。

4.1 预处理的步骤

上一环节，是通过图像检索技术，将当前图像在云数据库中的相似图像找出。如果检索结果正确，那么这个图像与当前图像应该具有相同或者非

常相似的拍摄物体或场景。但是，需要注意的是，我们检索得到这个相似图像的目的，与地标检索、目标识别等不同[18]；我们不仅仅需要高层语义级别的匹配，更需要在低级别的像素或者图像块层面进行匹配，从而使得参考图像块与当前图像块之间的残差值更小。通过观察我们发现，检索得到的图像与当期图像，由于可能在不同的光照强度、不同的角度、不同的位置拍摄，同时也有可能由于照相参数如焦距、曝光时间、光圈大小的不同，它们在细节纹理上存在较大差异。因而，为了下一步采用块匹配的方式预测的准确性，我们需要在编码之前对检索得到的图像进行预处理，以满足图像间预测进行块匹配的需求。

通过对比观察检索得到的相似图像与当前图像，我们发现它们之中的物体大小、位置和角度不尽相同，同时两幅图像的亮度具有一定差别（如图 1 中的原始图像与检索得到的图像）。所以，我们对检索得到的图像进行两个方面的预处理：图像间的几何形变及光照补偿[1]。我们用 $\boldsymbol{I}$ 表示当前图像，$\boldsymbol{I}'$表示检索得到的相似图像。

4.2 相似图像的几何形变

几何形变，是通过将一个图像进行平移、旋转、放缩等操作后，与另一幅图像中的物体进行对齐。我们利用当前图像与相似图像之间匹配的特征点进行计算。根据透视变换原理[19], [20]，每四对匹配的特征点就可以计算一个投影变换矩阵，这个矩阵表示了图像之间的旋转、平移、放缩等形变信息。由于两幅相似图像之间大都存在较多匹配的特征点，从而可以得到很多个变换矩阵。从中找出一个最佳变换矩阵的方法，最常用的是由 Fischler et al. 提出的 RANSAC 方法[21]，[22]。而在两个相似图像之间可能不仅存在一种变换，如果有多个物体或者一个物体有多个组成部分，那么两图像之间会存在多种变换。考虑到这种情况，我们不使用 RANSAC 方法，而是通过求解优化问题来选出最优的若干个变换矩阵。这个优化问题形式化表达成以下能量方程的求解

$$E = d + \eta \times s, \tag{1}$$

其中，d 是数据项；s 是平滑项；η 是权重系数，用来平衡两项的贡献量。

每四对特征点决定一个变换矩阵，这样由所有匹配的特征点对就得到多个变换矩阵，公式(1)的目标，就是从这些所有的变换矩阵中选出一个子集□，使得能量方程的值最小。在公式(1)中，数据项 d 表示相似图像中的所有特征点在变换集□下到当前图像中匹配特征点的距离之和，它的具体表达式如下

$$\begin{aligned} d &= \sum d_i, \\ d_i &= \left\| \boldsymbol{p}(\boldsymbol{f}_i) - \boldsymbol{p}(\boldsymbol{f}_i^{'}) \times \boldsymbol{H} \right\|. \end{aligned} \tag{2}$$

其中，$\boldsymbol{H}$ 代表集合□中的一个变换矩阵；$\boldsymbol{f}_i$ 和 $\boldsymbol{f}_i'$分别表示当前图像及其相似图像中的第 i 对匹配的特征点；$\boldsymbol{p}(*)$表示特征点的位置；d_i 则反映了这一对特征点在 $\boldsymbol{H}$ 变换下的距离。

在公式(1)中，平滑项 s 表示特征点之间的连通性，用点集的三角剖分（Delaunay triangulation）表征。s 的值，表示为每一对在三角剖分网格中相邻的特征点的平滑度 s_i 之和，即

$$\begin{aligned} s &= \sum s_i, \\ s_i(j,k) &= \begin{cases} 0, & \boldsymbol{p}(\boldsymbol{f}_j) \text{和} \boldsymbol{p}(\boldsymbol{f}_k) \text{同时满足变换} \boldsymbol{H} \\ 1, & \text{其它} \end{cases}. \end{aligned} \tag{3}$$

对公式 (1) 求解，我们通过图割（graph cut）的方法[22]，找到其最优的变换集 $\mathrm{H} = \{\boldsymbol{H}_1, \boldsymbol{H}_2, \cdots, \boldsymbol{H}_N\}$。其中，$N$ 为最优变换矩阵的个数，它的值一般比较小，决定用作预测的参考图像数量。然后，根据求解得到的这 N 个变换矩阵，对相似图像 $\boldsymbol{I}'$进行形变，相应地得到 N 个形变后的图像$\{\boldsymbol{I}^{\boldsymbol{H}_1}, \boldsymbol{I}^{\boldsymbol{H}_2}, \cdots, \boldsymbol{I}^{\boldsymbol{H}_N}\}$ 。

4.3 相似图像的光照补偿

上一步骤的几何形变，本质上通过移动相似图像上物体的像素点位置，将它与当前图像中的同一

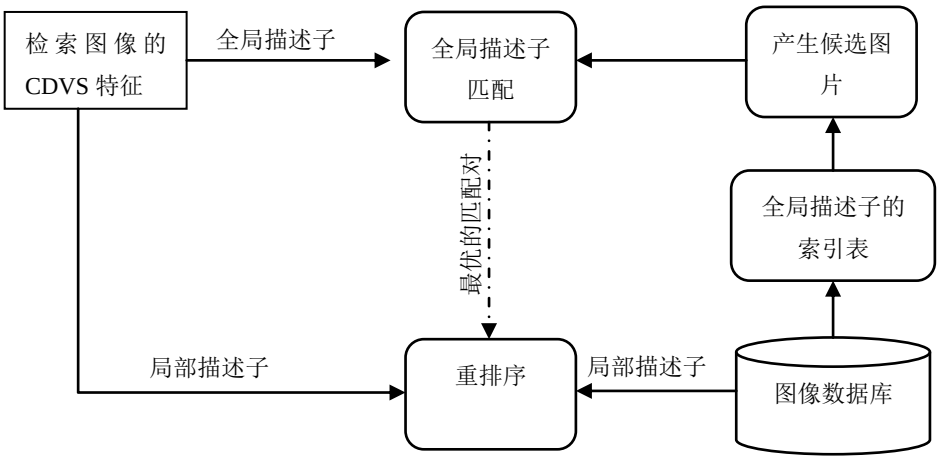


图 4 利用 CDVS 进行图像检索的流程图

物体在位置、大小、角度等方面做好了匹配。考虑到同一个物体上的对应像素点，在不同的光照强度下拍摄，像素值会有明显差异，所以我们继续通过光照补偿使得变换后的相似图像更为接近于当前图像。

相似图像 $\boldsymbol{I}'$ 经过变换 $\boldsymbol{H}_l, 1 \le l \le N$，形成了图像 $\boldsymbol{I}^{\boldsymbol{H}_l}$，我们对其进行线性亮度变换。这个过程用以下公式表达，

$$\boldsymbol{I}'' = \alpha \boldsymbol{I}^{\boldsymbol{H}_l} + \beta, \tag{4}$$

其中，α 和 β 分别表示尺度和偏移光照补偿参数，$\boldsymbol{I}''$ 是对图像 $\boldsymbol{I}^{\boldsymbol{H}_l}$ 经过光照补偿之后的形成的新图像。光照补偿的目标，是使得图像 $\boldsymbol{I}''$ 与当前图像 $\boldsymbol{I}$ 尽可能接近，这表示为下面的公式

$$\min \sum_{1 \le i \le M} \left(\boldsymbol{I}_{\boldsymbol{p}(\boldsymbol{f}_i)} - \boldsymbol{I}''_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')} \right). \tag{5}$$

其中，M 为所有匹配特征点的个数。$\boldsymbol{I}_{\boldsymbol{p}(\boldsymbol{f}_i)}$ 表示在当前图像中，位于特征点 $\boldsymbol{f}_i$ 位置 $\boldsymbol{p}(\boldsymbol{f}_i)$ 处的像素值；$\bar{\boldsymbol{p}}(\boldsymbol{f}_i')$，表示相似图像 $\boldsymbol{I}'$ 中与 $\boldsymbol{f}_i$ 匹配的特征点 $\boldsymbol{f}_i'$ 经过变换 $\boldsymbol{H}_l$ 后，在变换后的图像 $\boldsymbol{I}^{\boldsymbol{H}_l}$ 中的坐标位置；$\boldsymbol{I}''_{\boldsymbol{p}(\boldsymbol{f}_i)}$ 表示在经光照补偿后的图像 $\boldsymbol{I}''$ 中，位于位置 $\bar{\boldsymbol{p}}(\boldsymbol{f}_i')$ 处的像素值。 计算的目标是，使得每一对匹配特征点位置处的像素值差异尽可能小。结合公式(4)和公式(5)，我们便可得到以下方程

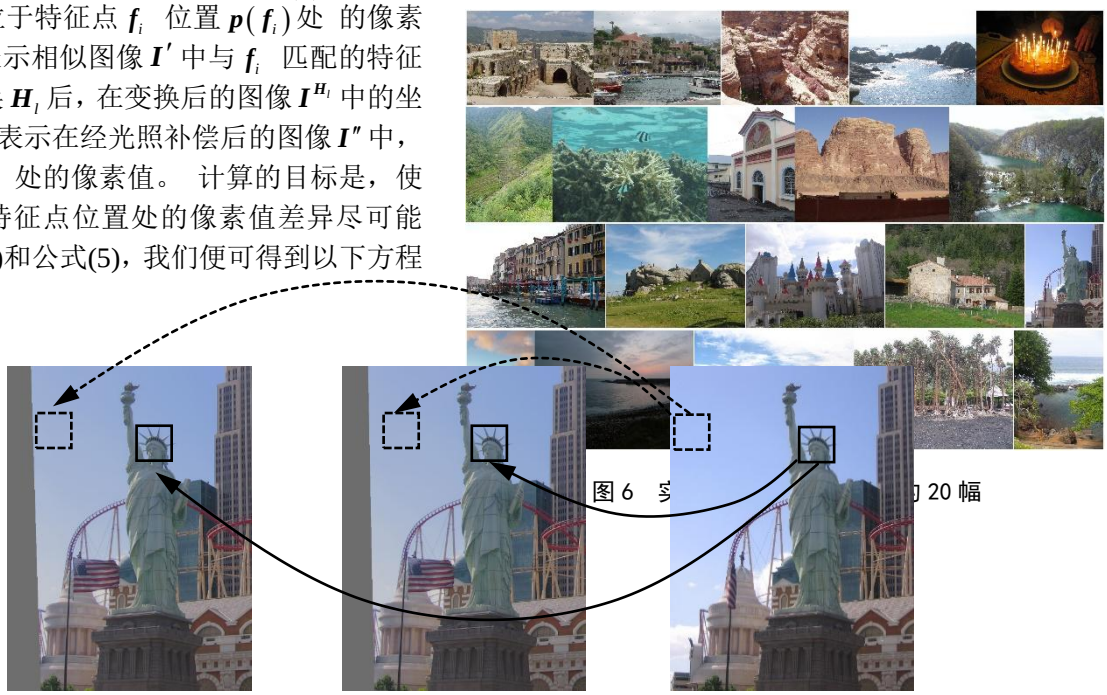


图 5　当前图像利用参考图像进行基于块匹配的预测编码

来估计光照补偿的参数 α 和 β

$$(\alpha, \beta) = \arg\min_{\alpha,\beta} \left(\sum_{1 \le i \le M} \boldsymbol{I}_{\boldsymbol{p}(\boldsymbol{f}_i)} - \left(\alpha \boldsymbol{I}^{\boldsymbol{H}_l}_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')} + \beta \right) \right). \tag{6}$$

利用最小二乘法[23]，我们可以得到参数 α 和 β 的解析表达式如下

$$\alpha = \frac{M \cdot \sum \boldsymbol{I}_{\boldsymbol{p}(\boldsymbol{f}_i)} \cdot \boldsymbol{I}^{\boldsymbol{H}_l}_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')} - \sum \boldsymbol{I}_{\boldsymbol{p}(\boldsymbol{f}_i)} \cdot \sum \boldsymbol{I}^{\boldsymbol{H}_l}_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')}}{M \cdot \sum \left(\boldsymbol{I}^{\boldsymbol{H}_l}_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')} \right)^2 - \left(\sum \boldsymbol{I}^{\boldsymbol{H}_l}_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')} \right)^2},$$

$$\beta = \frac{\sum \boldsymbol{I}_{\boldsymbol{p}(\boldsymbol{f}_i)} \cdot \sum \left(\boldsymbol{I}^{\boldsymbol{H}_l}_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')} \right)^2 - \left(\sum \boldsymbol{I}_{\boldsymbol{p}(\boldsymbol{f}_i)} \cdot \boldsymbol{I}^{\boldsymbol{H}_l}_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')} \right) \cdot \sum \boldsymbol{I}^{\boldsymbol{H}_l}_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')}}{M \cdot \sum \left(\boldsymbol{I}^{\boldsymbol{H}_l}_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')} \right)^2 - \left(\sum \boldsymbol{I}^{\boldsymbol{H}_l}_{\bar{\boldsymbol{p}}(\boldsymbol{f}_i')} \right)^2}. \tag{7}$$

根据公式(7)，对于变换 $\boldsymbol{H}_l$，我们可以求解得到公式(4)中的参数 α 和 β；然后根据公式(4)，就可得到图像 $\boldsymbol{I}^{\boldsymbol{H}_l}$ 经过光照补偿后的图像，用 $\boldsymbol{I}^{P_l}$ 表示。如果在上一步骤的几何形变中，得到了多个变换矩阵，即 N>1，则这一步的光照补偿也可对应得到 N 对光照参数，最后可得 N 个补偿后的图像。

这样，经过对相似图像的两步预处理过程，我们可以得到 N 个图像作为当前待编码图像的参考图像（图 1 中以 2 个参考图像作为示例）。

5 图像间预测编码

在本章中，我们对整个云图像编码系统中的最后一个环节----图像间的预测编码进行介绍。首先介绍基于图像块匹配的预测编码；然后介绍基于率失

真优化的模式决策。

5.1 基于图像块匹配的预测编码

在上一环节对相似图像进行预处理之后，得到的 N 个图像与当前图像更为接近，但并不一定完全一样。一部分原因是，当前图像存在有相似图像中不存在的图像内容（比如，图 1 中当前图像比变换后的相似图像多了左侧的蓝色部分）；还有一部分原因是，预处理中的几何形变和光照补偿估计的参数可能存在误差，或者它不能满足整幅图像中所有内容的要求。因此，我们对于由预处理得到的参考图像，并不是整体作为当前图像的预测，而是借用视频编码中运动估计的思想[24]，[25]，先以图像块为单位进行块匹配。

HEVC (High Efficiency Video Coding) [26], [27] 是国际上最新一代视频编码标准，它采用层级递归的块划分方式，依此进行模式选择和运动估计。在我们这一步的编码中，首先按照 HEVC 的方式对当前待编码图像进行块划分。对于每一个图像块，如图 5 所示，我们在那 N 个参考图像中搜索其相似块（图 5 中的参考图像个数为 2）。相似块的标准为与当前图像块的均方误差（mean square error, MSE）最小的图像块，用数学形式表达如下

$$Sim(\boldsymbol{B}) = \arg\min_{\boldsymbol{B}_i} \left\| \boldsymbol{B}_i - \boldsymbol{B} \right\|_2^2 . \tag{8}$$

其中，$\boldsymbol{B}$ 表示当前图像块，$\boldsymbol{B}_i$ 是参考图像中的候选图像块，$Sim(\boldsymbol{B})$表示当前图像的相似块。如图 5 所示，按照这种方式，实线方框里自由女神的头部可以在两个参考图像中都找到非常接近的块，选均方误差最小的作为这个块的预测。而虚线方框区域实际上在参考图像中并没有完全对应的图像内容，但是通过相似块查找，我们依然可以在参考图像中找到与其足够接近的图像块，如其它位置的蓝天部分。通过记录这个块所在的参考图像以及它与当前块的位置偏差，用这个块作为当前块的预测进行编码。

5.2 基于率失真优化的模式决策

上一部分通过块匹配确定预测块的方式，是假定参考图像和当前图像足够相似，并且对于当前块，有在灰度值上足够接近的块可以提供预测。而实际系统中也可能存在这样的情况，即参考图像跟当前图像并非包含同样的物体，或者参考图像中的块不能为当前图像块提供足够好的预测（尽管参考图像已经经过了预处理）。在这种情况下，如果仍然用找到的“相似块”进行预测，那么它与当前块的残差将会很大，这样编码产生的比特数并不一定比单独编码当前图像要少，甚至产生的额外开销可能会大于由图像间预测带来的节省。所以，我们需要一种自适应的机制，使得在参考图像不能提供有效预测块的情况下，可用另一种方式自动替代图像间相似块查找的预测方式，从而保证系统仍然能够高效编码。

率失真优化（rate-distortion optimization, RDO）技术[28]，是视频编码标准中广泛被采用的决策编码模式的方式。它的基本思想是，对于一个待编码图像块，用不同的模式编码会产生不同的率失真代价（rate-distortion cost, RD Cost），我们通过率失代价如公式(9)

$$J = D(m_i) + \lambda R(m_i) \tag{9}$$

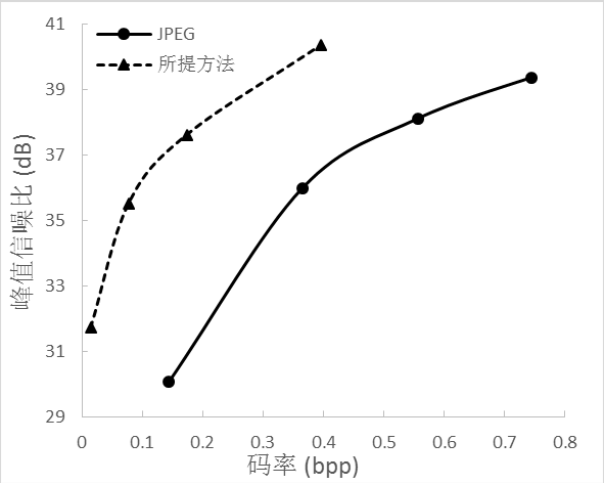


图 7　本文算法与 JPEG 平均率失真性能比较

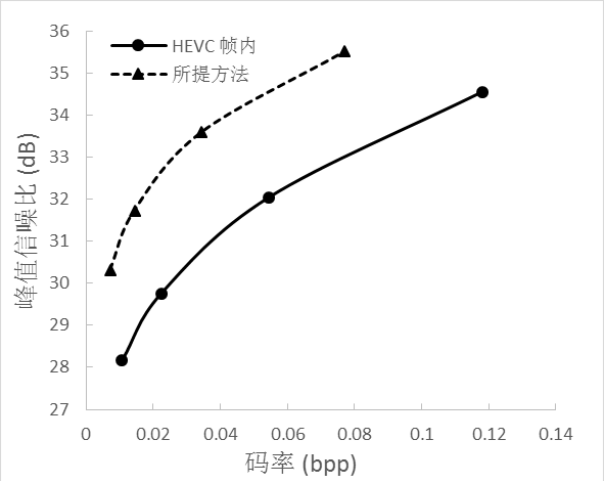


图 8　本文算法与 HEVC 帧内平均率失真性能比较

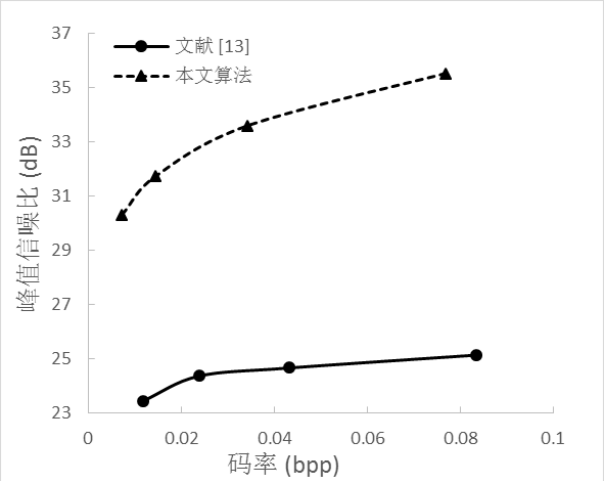
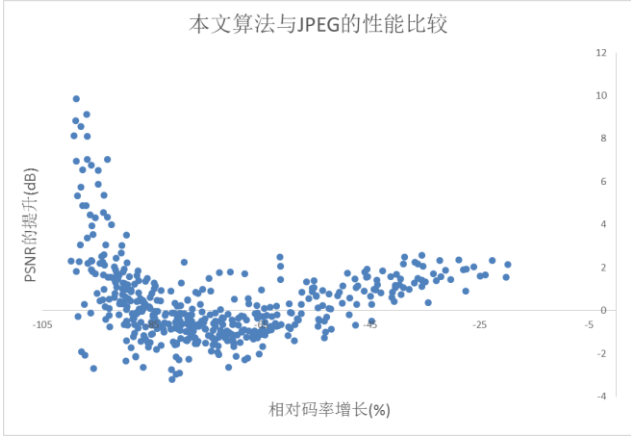
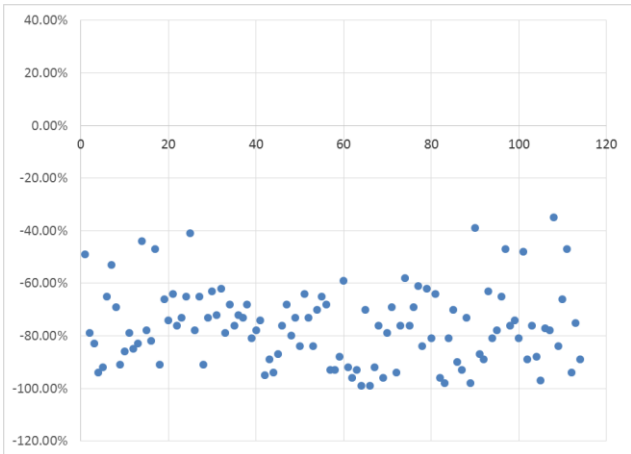


图 9　本文算法与文献[13]平均率失真性能比较



(a)



(b)

图 10　(a) 对每幅测试图像，本文算法相对于 JPEG 的 PSNR 值的提升与相对码率增长的关系图；(b) 对每幅图像，本文算法相对于 JPEG 的 bdrate

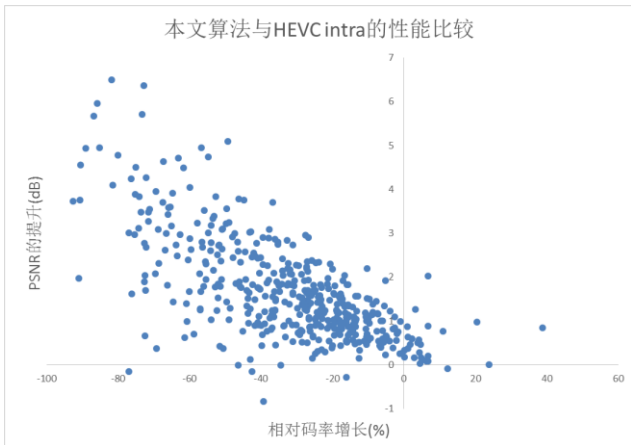


图 11　对每幅测试图像，本文算法相对于 HEVC intra 的 PSNR 值的提升与相对码率增长的关系图

的计算，选择使其最小的模式进行图像块的编码。在公式(9)中，m_i 表示某一种编码模式；$D(m_i)$ 表示在这种模式下编码当前块所产生的误差，一般通过当前块的重构值与原始值的差的平方和（sum of square difference, SSD）来计算；$R(m_i)$ 表示用这种模式编码当前块所产生的码流比特数；λ 是拉格朗日乘子，用来协调两项所占的比重，由量化参数（quantization parameter, QP）确定。模式决策的过程可以用如下公式表示

$$\text{mode}_{opt} = \arg\min_{m_i} J, \tag{10}$$

其中，mode_{opt} 是选出来的最优编码模式。编码模式分为图像间编码和图像内部编码，在第 5.1 中介绍的基于块匹配的预测编码是图像间的预测编码方案，图像内部的编码我们参照 HEVC [26]帧内预测方式进行。在图像间编码方法中，我们编码预测块与当期块的残差，以及标识预测块的参考图像索引和位置差等信息，计算所需的比特数，得到公式(9)中的 R；然后对码流信息进行解码，得到这种方式下重构的图像块，计算它与当前块的原始值的差异，得到公式(9)中的 D。对于图像内部编码，率失真代价的计算采取类似的方式。这样，通过如公式(10)所示的模式决策过程，系统判断图像间编码和图像内部编码的率失真代价大小，从中选择代价较小者作为实际的编码方式。

根据上述过程，如果在云图像数据库中不存在当前图像的相似图像，那么对当前图像的编码可以自动切换到图像内部编码方式。这种机制，保证了图像的基本重建质量，避免了[13]中一旦没有相似图像当前图像无法正确重构的问题。

在确定了编码模式后，系统对残差通过变换、量化和熵编码进行后续编码。

6 实验结果与分析

为了进一步验证本文所提出图像编码算法的有效性，我们在“INRIA Holidays”图像数据集[3]上进行了大量实验。“INRIA Holidays”图像数据集，是由“INRIA LEAR” 项目组和“Advestigo”公司为“ANR RAFFUT” 项目联合创建，总共有 1491 幅图片。它里面的图片大多为世界各地的景点照片，每个景点有二到十几幅在不同角度、焦距、光照等条件下拍摄的图片，其中有一幅选作“query”图片，其它的作为“database”图片。整个数据集有 500 幅“query”图片和 991 幅“database”图片。文献[13]的实验测试，也是基于这个数据集。在实验中，我们从“query”图像中选择了 114 幅作为测试图像，将其它图像作为云数据库中的图像。在图 6 中，我们展示了其中的 20 幅测试图像。

对所有测试图像，分别用本文所提出的基于云数据的图像编码算法、JPEG、HEVC 帧内编码（当前图像设为 HEVC 中的 I 帧）以及文献[13]的方法进行压缩和重建。JPEG 编码，采用 matlab2012a 中提供的方法；HEVC 帧内，采用 HM 18.4 中的“all intra”配置；对于文献[13]的方法的测试，是基于该文献作者提供的源代码。

图 7 展示了本文所提出算法和 JPEG 对测试图像的平均率失真性能对比，图 8 展示了本文算法和 HEVC 帧内对测试图像的平均率失真性能对比。图 9 展示了本文算法和文献[13]的方法对测试图像的平均率失真性能对比。在图 7、图 8 和图 9 中，曲线的横坐标表示平均编码每个像素点所用的比特数（单位为 bpp），纵坐标表示重建图像的峰值信噪比（peak signal to noise ratio, PSNR，单位为 dB）。在横坐标值固定的情况下，PSNR 值越高，表示编码性能越好。从图 7 和图 8 中可以看出，本文算法的压缩性能明显高于传统的 JPEG 图像编码方法；而且相比于最新一代视频编码标准 HEVC 的高效帧内编码方式，本文算法也有明显的性能提升。通过 bdrate[29]的计算得出，提出的方法跟 JPEG 相比，平均码率节省 78.5%；跟 HEVC 帧内编码方式相比，平均码率节省 67.2%。

从图 9 中可以看出，本文算法相比于文献[13]的方法，在客观质量的保真度上有非常明显的优势。在同样码率的情况下，本文方法的平均 PSNR 值明显高于[13]重建图像的平均 PSNR。需要说明的是，本文提供的文献[13]的实验结果，与其论文中给出的结果一致（PSNR 在 20 dB 上下）：文献[13]中的数据也表明，其方法的客观质量的分值，甚至比传统的 JPEG 都要低。由于本文算法在客观质量 PSNR 数值上比文献[13]的方法高出很多，已经超出了 bdrate 能够正确衡量性能的范围，所以我们没有给出具体增益的数值。在下文中，我们会通过图 12 给出每个图像的性能对比，直观地表现出我们算法

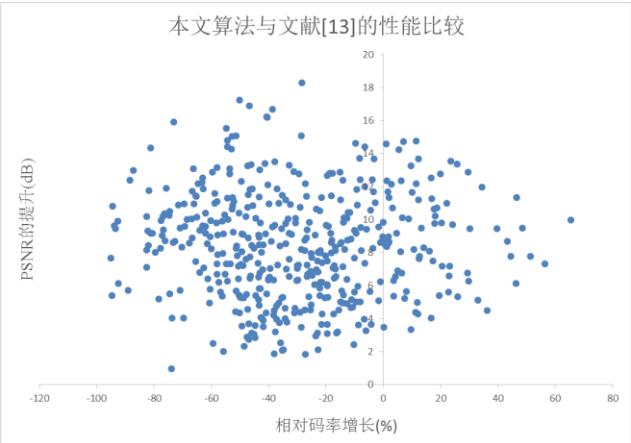


图 12 对每幅测试图像，本文算法相对于方法[13]的 PSNR 值的提升与相对码率增长的关系图

3 http://lear.inrialpes.fr/~jegou/data.php

的性能增幅。

图 10、图 11 和图 12 中，展示了每一幅测试图像的实验结果，其中每一个点代表一幅图像。图 10（a）表示了，本文算法相对于 JPEG 的 PSNR 提升和相对码率增长的关系；图 11 表示了，本文算法相对于 HEVC intra 的 PSNR 提升和相对码率增长的关系；图 12 表示了，本文算法相对于文献[13]的方法的 PSNR 提升和相对码率增长的关系。根据图像视频编码的经验，A 方法跟 B 方法比较，如果在码率增长为 20%的情况下，PSNR 提升 1dB，那么 A 和 B 方法具有相当的性能；如果 A 方法的 PSNR 的增长明显超过了 1dB，或者码率的增长明显小于 20%甚至为负值，那么 A 方法相对于 B 方法有性能增益；反之，A 方法相对于 B 方法有性能损失。从图 10（a）中可以看出，本文算法对于所有测试图像，都有很大幅度的码率节省，而且在这种情况下大部分图像的 PSNR 还都有提升。而少部分在码率节省的同时 PSNR 下降的图像，图 10（a）不方便比较算法性能，于是我们同时给出了图 10（b）。图 10（b）中展示了，编码每幅图像用本文算法相对于 JPEG 方法的 bdrate 值。可以看出所有图像的 bdrate 均为负值，表示都有增益，而且大部分图像都有比较大的性能增益。在图 11 中，大部分点位于第二象限，这意味着本文算法相对于 HEVC intra 在有码率节省的情况下，还有客观质量的提升，具有非常明显的性能增益。在图 12 中，同样大部分点位于第二象限；同时，少数位于第一象限的点，相对于码率增长的数值，PSNR 提升的数值远远超过了性能持平的值。因此，图 12 也反映出了本文算法相比于文献[13]的方法，对绝大多数测试图像都有客观质量的提升。

图 13、图 14 和图 15 展示了，用各个方法编码并重建得到的图像的主观质量对比。我们让各个方法的码率尽可能接近，来对比解码图像的质量。本

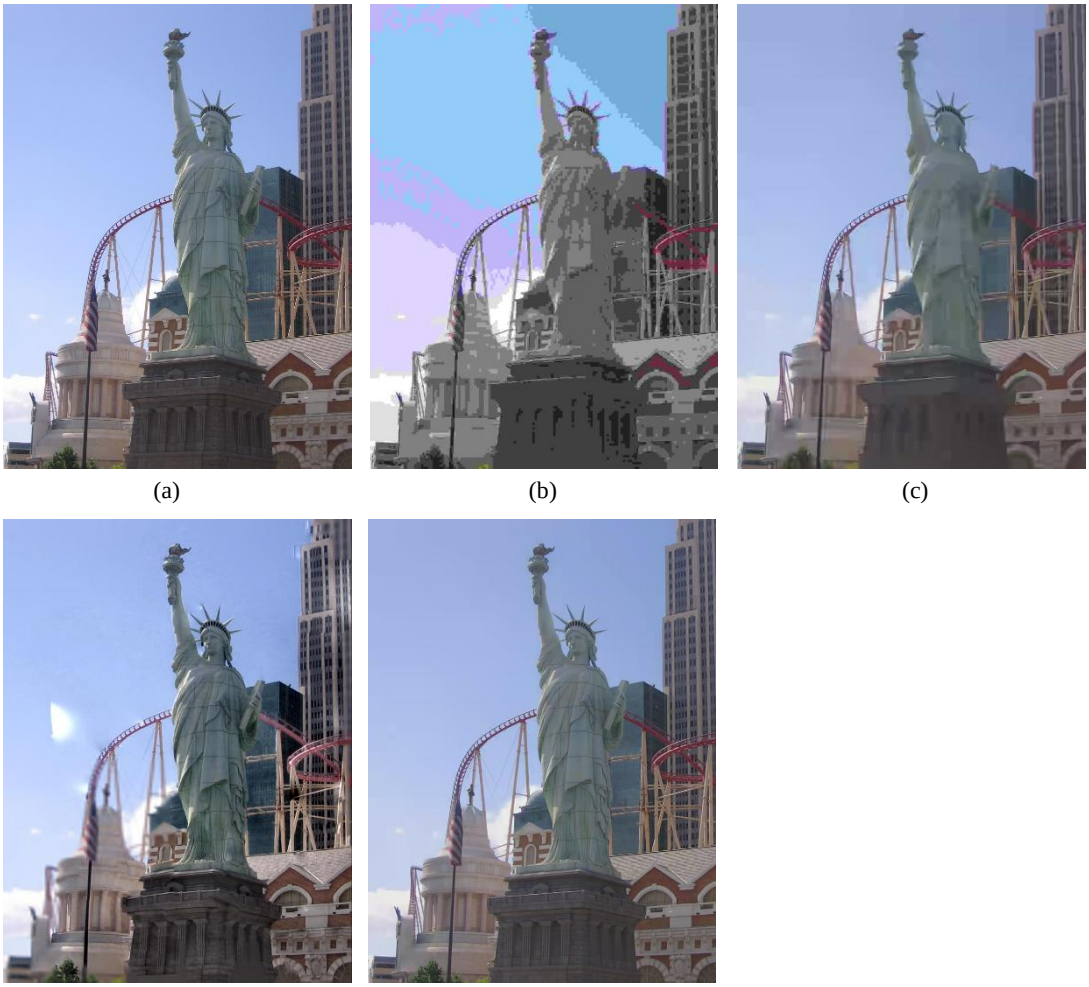


图 13 本文算法与 JPEG、HEVC 帧内和文献[13]的方法解码图像的主观质量比较。(a)原始图像；(b)JPEG 的解码图像，bpp=0.145，PSNR=29.78dB；(c)HEVC 帧内的解码图像，bpp=0.019， PSNR=30.64dB；(d)文献[13]方法的解码图像，bpp=0.020，PSNR=24.24dB；(e)本文算法的解码图像，bpp=0.016，PSNR=35.45dB.

文算法、HEVC 帧内和文献[13]的方法都采用 bpp=0.01 码率级；由于 JPEG 在 bpp=0.1 码率级质

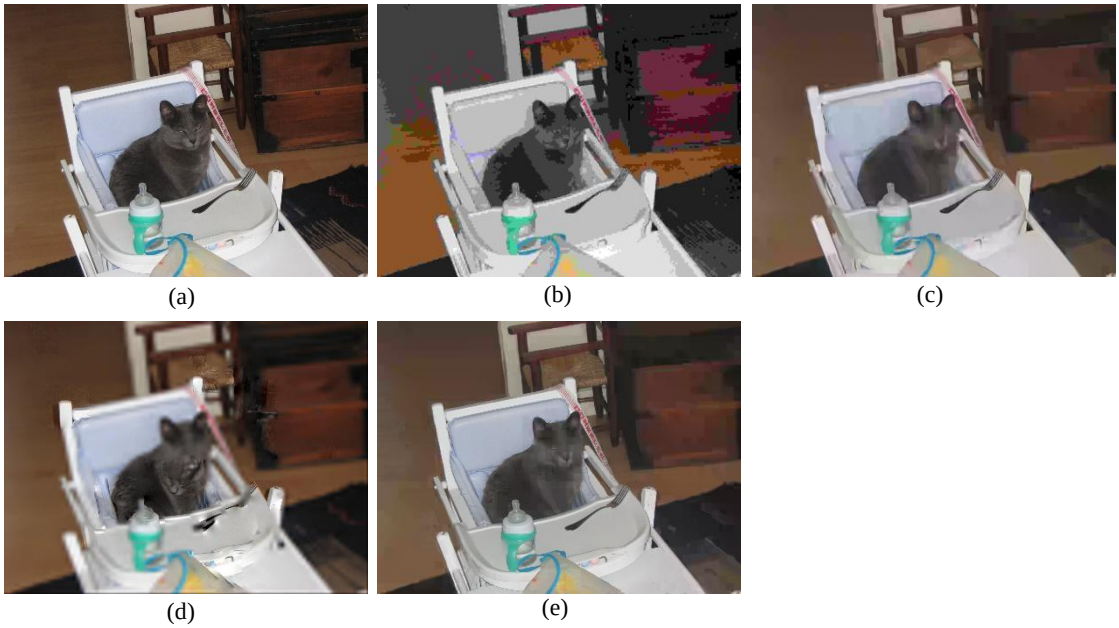


图 14 本文算法与 JPEG、HEVC 帧内和文献[13]的方法解码图像的主观质量比较。(a)原始图像；(b)JPEG 的解码图像，bpp=0.139，PSNR=30.33dB；(c)HEVC 帧内的解码图像，bpp=0.008， PSNR= 30.93dB；(d)文献[13]方法的解码图像，bpp=0.012，PSNR=20.57dB；(e)本文算法的解码图像，bpp=0.006，PSNR=32.86dB。

图 15 本文算法与 JPEG、HEVC 帧内和文献[13]的方法解码图像的主观质量比较。(a)原始图像；(b)JPEG 的解码图像，bpp=0.421，PSNR=30.66dB；(c)HEVC 帧内的解码图像，bpp=0.015， PSNR=30.56dB；(d)文献[13]方法的解码图像，bpp=0.018，PSNR=27.15dB；(e)本文算法的解码图像，bpp=0.014，PSNR=31.68dB.

量已经很低，没有必要继续压缩码率。从图中可以看出，本文算法不仅在客观质量（PSNR）上取得最高的分数，而且重建图像的主观质量对于原始图像也有最高的保真度。而相比之下，JPEG 解码图像有明显的块效应，而 HEVC 帧内方法的解码图像丢失了大量的纹理细节。文献[13]方法的重建图像，在主观质量上存在区域性差异。对于能够在库中找到相似块的区域，它能够恢复出较为清晰的图像，如图 13 (d)中的雕像部分；在不能够检索到相似块的区域，重建质量就依赖于上采样后的解码小图像，视觉效果就非常模糊，如图 13(d)旗子及其左侧部分。同时，由于不能保证图像块形变的准确性以及将图像块到整幅图的完美拼接，重建图像中会出现一些线条的扭曲以及接合处的错位，如旗杆后方建筑柱子的扭曲以及“自由女神”右脚前方的错位。相比之下，本文算法并不依赖于云中是否存在相似图像以及形变的准确性。它在构建完参考图像后，进行了以块为单位的相似块搜索，从更细粒度上去做匹配；同时，它结合了视频编码中的率失真优化技术，当参考图像不能提供准确的预测时，还能够采用图像内编码的方式。可以说，它不仅利用了图像间预测编码的高效性，同时也保留了图像内编码来保证最坏情况下的基本重建质量。

7 总结

本文提出了一种基于云数据的图像编码方法。该方法通过利用云数据库中已有的图像数据，充分挖掘图像间的相关性进行编码，大大提高了压缩效率。该算法结合图像检索技术、图像预处理技术以及图像间预测编码技术，将对单幅图像的编码转换为对其残差的编码，从而减少了编码所产生的数据量。同时，算法提供了图像内编码补偿机制，能够保证在图像库中不存在相似图像时的图像重建质量。

该算法为大数据时代的图像编码提供了一个崭新的框架，可以应用于如网络云盘、社交网络图像等的图像压缩。基于该算法框架，可以进一步研究云图像编码算法。下一步的研究工作包括：云图像库的组织方式，更精确的图像检索算法，优化参考图像预处理等。

致 谢 本文作者对文献[13]的作者表示衷心的感谢，感谢他们提供其文章的源代码，便于本文算法与其算法性能的比较。同时，还要感谢“INRIA Holidays”图像数据集[4]原创者们的分享，能让我们更方便地测试算法性能。

参 考 文 献

[1] Z. Shi, X. Sun and F. Wu. Feature based Image Set Compression. Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), San Jose, USA, 2013, 1-6

[2] Shi, Zhongbo, Xiaoyan Sun, and Feng Wu. Multi-model prediction for image set compression. Proceedings of the IEEE International Conference on Visual Communications and Image Processing (VCIP). Kuching, Malaysia, 2013: 1-6.

[3] Kosmas Karadimitriou and John M. Tyler. The centroid method for compressing sets of similar images. Pattern Recognition Letters, 1998, 19(7): 585 – 593.

[4] Yurij S. Musatenko and Vitalij N. Kurashov. Correlated image set compression system based on new fast efficient algorithm of karhunen-loeve transform. Photonics East, 1988, 3527(10): 518–529.

[5] Samy Ait-Aoudia and Abdelhalim Gabis. A comparison of set redundancy compression techniques. EURASIP Journal on Applied Signal Processing, 2006(15): 1–13.

[6] Cheng H. Li X. Schmieder, A. A study of clustering algorithms and validity for lossy image set compression. Proceedings of the International Conference on Image Processing, Computer Vision and Pattern Recognition (IPCV). Las Vegas, Nevada, USA, 2009: 501 –506.

[7] A. George, E. Yasser, H.Meer. Quadtree-Based Centroid Technique for Compressing Sets of Similar Medical Images. Proceedings of the 4th International Conference on Intelligent Systems Design and Applications. Budapest, Hungary, 2004: 1-6.

[8] Chi-Ho Yeung, Oscar C. Au, Ketan Tang, Zhiding Yu, Enming Luo, Yannan Wu, and Shing Fat Tu. Compressing similar image sets using low frequency template. Proceedings of the IEEE International Conference on Multimedia and Expo (ICME). Barcelona, Spain, 2011: 1 –6.

[9] C. Nielsen and Xiobo Li. MST for lossy compression of image sets. Proceedings of the Data Compression Conference. Snowbird, Utah,

[4] http://lear.inrialpes.fr/~jegou/data.php

US, 2006: 463 –463.

[10] B. Gergel, H. Cheng, C. Nielsen, and X. Li. A unified framework for image set compression. Proceedings of the International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV), Las Vegas, Nevada, USA, 2006: 417–423.

[11] B. Gergel, A. Schmieder, X. Li, and H. Cheng. A study of prediction measures for lossy image set compression. Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV). Las Vegas, Nevada, USA, 2008: 69–74.

[12] J. Ranger and H. Cheng. Structural similarity as a prediction metric in lossy image set compression. Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV). Las Vegas, Nevada, USA, 2012:1-6.

[13] Huanjing Yue, Xiaoyan Sun, Jingyu Yang, and Feng Wu. Cloud-Based Image Coding for Mobile Devices—Toward Thousands to One Compression. IEEE Transactions on Multimedia, 2013, 15(4): 845-857.

[14] Ru Li-Yun, Peng Xiao, Su Zhong, and Ma Shao-Ping. Feature performance evaluation in content-based image retrieval. Journal of Computer Research and Development, 2004, 40(11): 1566-1570 (in Chinese)
(茹立云，彭潇，苏中，等. 基于内容图像检索中的特征性能评价. 计算机研究与发展, 2004, 40(11): 1566-1570)

[15] Guo Li, Huang Yuanyuan, and Yang Jingyu. Using sub-block image features to retrieve trademark image. Journal of Computer Aided Design & Computer Graphics, 2004, 16(7): 968-972 (in Chinese)
(郭丽，黄元元，杨静宇. 用分块图像特征进行商标图像检索. 计算机辅助设计与图形学学报, 2004, 16(7): 968-972)

[16] Li Bing, Liu Lei, Wei Zhi-Qiang. A strong robust real-time image matching algorithm. Journal of Software, 2014, 25(7): 1583-1592 (in Chinese)
(李兵，刘磊，魏志强. 一种具有强实时性，强鲁棒性的图像匹配算法. 软件学报, 2014, 25(7): 1583-1592)

[17] Ling-Yu Duan, Jie Lin, Jie Chen, Tiejun Huang, Wen Gao. Compact Descriptors for Visual Search. IEEE Multimedia, 2014, 21(3): 30-40.

[18] Song Xiao-Xuan, Liu Jing-Hao, Tang Xiao-Rong. Image Retrieval-based Landmark Recognition System. Electornic Design Engineering, 2012, 20(12): 173-175 (in Chinese)
（宋小璇，刘敬浩，唐小容. 基于图像检索的地标识别系统. 电子设计工程, 2012, 20(12):173-175.）

[19] R. Szeliski. Computer vision: algorithms and applications. Berlin, Germany: Springer Science & Business Media, 2010.

[20] Zeng Wen-Feng, Li Shu-Shan, Wang Jiang-An. Translation, rotation and scaling changes in image registration based affine transformation model. Infrared and Laser Engineering, 2001, 30(1):18-20 (in Chinese)
(曾文锋，李树山，王江安. 基于仿射变换模型的图像配准中的平移、旋转和缩放. 红外与激光工程, 2001, 30(1):18-20)

[21] Shan Xin, Wang Yao-Ming, Dong Jian-Ping. The Matching Method Based on RANSAC Algorithm for Estimation of the Fundamental Matrix. Qu Tian-Wei, An Bo, Chen Gui-Lan. Application of Improved RANSAC algorithm to Image Registration. Journal of Computer Applications, 2010, 7(7): 1849-1851 (in Chinese)
(曲天伟，安波，陈桂兰. 改进的RANSAC算法在图像配准中的应用. 计算机应用, 2010, 7(7): 1849-1851)
Fischler, M.A. and Bolles, R.C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Communications of the ACM, 24(6): 381–395, 1981.

[22] Zou Kun, Han Guo-Qiang, Li Wen, Zhang Xiao-Yuan. An efficient Method of Texture Synthesis Based on Graph Cuts. Journal of Computer-Aided Design and Computer Graphics, 2008, 20(5),: 652-658 (in Chinese)
（邹昆，韩国强，李闻，张潇元. 基于 Graph Cut 的快速纹理合成算法. 计算机辅助设计与图形学学报, 2008, 20(5): 652-658）

[23] Zhang Hao-Ran, Wang Xiao-Dong. Incremental and Online Learning Algorithm for Regression Least Squares Support Vector Machine, 2006, 29(3): 400-406 (in Chinese)
(张浩然，汪晓东. 回归最小二乘支持向量机的增量和在线式学习算法. 计算机学报, 2006, 29(3): 400-406.)

[24] Wu Xiao-Jun, Bai Shi-Jun, Lu Wen-Tao. Optimization on Motion Estimation Algorithm Based on H.264. Acta Elecronica Sinica, 2009, 11(11): 2541-2545 (in Chinese)
(吴晓军，白世军，卢文涛. 基于 H.264 视频编码的运动估计算法优化[J]. 电子学报, 2009, 11(11): 2541-2545)

[25] Gao Wen, Wang Qiang, Ma Si-Wei. Digital Audio Video Coding Standard of AVS. ZTE Communications, 2006,3: 6-9 (in Chinese)
(高文，王强，马思伟. AVS 数字音视频编解码标准. 中兴通讯技术, 2006, 3: 6-9)

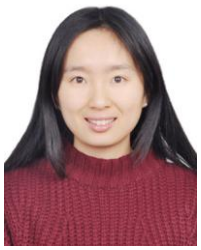
[26] Zhu Xiu-Chang, Li Xin, Chen Jie. Next generation video coding standard—HEVC. Journal of Nanjing University of Posts and Telecommunications(Natural Science), 2013, 33(3): 1-11 (in Chinese)
(朱秀昌，李欣，陈杰. 新一代视频编码标准——HEVC. 南京邮电大学学报: 自然科学版, 2013, 33(3): 1-11)

[27] Shen Yan-Fei, Li Jin-Tao, Zhu Zhen-Min, Zhang Yong-Dong. High efficiency video coding. Chinese Journal of Computers, 2013, 36(11):

2340-2355 (in Chinese)
(沈燕飞, 李锦涛, 朱珍民, 等. 高效视频编码. 计算机学报, 2013, 36(11): 2340-2355)

[28] Ma Si-Wei. Rate Distortion Optimization Based Video Coding[Doctoral dissertation]. Institute of Computing Technology Chinese Academy of Sciences, Beijing, 2007 (in Chinese)
(马思伟. 基于率失真优化的视频编码研究[博士学位论文]. 中国科学院计算机技术研究所，北京，2007)

[29] G. Bjontegaard. Calculation of average PSNR difference between RD-curves. Proc. ITU-T Q.6/SG16 VCEG 13th Meeting. Austin, TX, US, 2001: 2-4.



ZHAO Chen, born in 1988, Ph. D. candidate. Her research interests include image/video processing, image/video compression.

MA Si-Wei, born in 1979, Ph. D., professor. His main research interests include image and video coding, video processing, video streaming and transmission.

ZHANG Xin-Feng, born in 1983, Ph. D. His main research interests include image/video processing, image/video compression.

ZHANG Jian, born in 1985, Ph. D. His main research interests include image and video compression, processing.

GAO Wen, born in 1956, Ph. D., professor,Member of Chinese Academy of Sciences. His main research interests include image processing, video coding and communication, pattern recognition, multimedia information retrieval, multimodal interfaces, etc.

Background

The efforts in image compression have been lasting for decades of years. Conventional ways of image compression are removing redundancy inside an image by transformation, quantization and entropy coding. Nowadays, due to the rapid spread of personal smart phones, the number of images is growing in an unprecedented high speed. Conventional image compression methods can no more satisfy the expansion of image data. New methods are to be explored to much more efficiently compress, store or transmit the image data.

Cloud computing is becoming increasingly popular nowadays. With its development, a lot of cloud services (such as cloud storage) are emerging. We are trying to solve the problem of image compression in the background of cloud computing and by utilizing the benefit of cloud data. Although image compression is an old topic in the signal processing field, combining it with cloud data to reach an even higher compression ratio is a new research focus. There are not too many works in this area, and the techniques and systems in the literature are not mature yet. Ensuring the reconstruction quality in the absence of similar images in the cloud, as well as achieving high objective reconstruction quality is still an issue. We solve these two key issues in this paper, proposing a mechanism to use intra-image coding as a compensation scheme. The proposed system has significant compression performance gains compared to traditional image coding methods (e.g., JPEG) and even the advanced HEVC intra coding. Compared to the existing state-of-the-art attempt for cloud-based image coding, both the objective and subjective reconstruction quality of our algorithm has obvious improvement.

Our research group has devoted a lot of effort in high-efficiency image/video coding. We have a lot of papers published in the respectable international journals and conferences, such as T-CSVT, T-IP, ICIP, and DCC. Many of our proposals have been adopted in the video coding standards, such as H.264/AVC, HEVC and AVS. We are also making this new attempt to find novel framework or solutions for compressing image/video data. We have had a few paper published in this novel area, and have proposed our idea in the recent AVS meeting.

This work is supported in part by the National Science Foundation of China under Grants No.61322106 and No. 61370114, and by National Basic Research Program of China (973 Program) under Grant No. 2015CB351800. These projects aim to make advances for novel paradigm of image/video coding and reconstruction by utilizing methods and techniques different from traditional ways.