THE JPEG-BLOCKCHAIN FRAMEWORK FOR GLAM SERVICES

Deepayan Bhowmik, Ambarish Natu, Takaaki Ishikawa, Tian Feng and Charith Abhayaratne

Department of Computing, Sheffield Hallam University, Sheffield, United Kingdom, S1 1WB
Australian Government, Australian Capital Territory, Australia
Waseda University, Shinjuku-ku, Tokyo, Japan, 169-0072
Electronic & Electrical Engineering, The University of Sheffield, Sheffield, United Kingdom, S1 4DE
d.bhowmik@ieee.org, ambarish.natu@gmail.com, takaxp@ieee.org,
ashbringerft@gmail.com, c.abhayaratne@sheffield.ac.uk

ABSTRACT

This paper proposes a JPEG-blockchain framework for trusted media transaction. The new distributed and tamper-proof framework intends to aid an emerging JPEG Privacy and Security standard. The blockchain network records any media transaction with necessary information related to intellectual property rights, access control rules and content signature. The content signature, generated by compressed sensed samples or low-resolution, low bit-rate compression is used to verify the image integrity and authenticity. We propose that every blockchian record, linked to a unique *transaction hash*, is encapsulated within the metadata contained in the JPEG box structure. As an example use case we have chosen the GLAM (Galleries, Libraries, Archives and Museums) sector due to its emerging need. This paper presents the proof of the concept and reports preliminary infrastructural development.

***Index Terms*—** JPEG Privacy and Security, Blockchain, GLAM, Compressive sensing, JPEG-Blockchain.

1 Introduction

The GLAM (Galleries, Libraries, Archives and Museums) sector deals with a huge number of digital images and confronted with issues including IPR (Intellectual Property Rights) for access and usage control; authenticity (fake images and tampering); and information related to transaction histories [1]. A blockchain framework is proposed here to address such issues. The proposed framework ensures tamper-proof and authentic image transaction (*e.g.*, loan to other GLAM for exhibition or copyright transfer) enriched with transaction trails [2]. This work also intends to compliment the work on-going within the JPEG Privacy and Security framework [3].

A functional overview of our JPEG-blockchain framework is shown in Fig. 1. We propose that any JPEG image file that needs its transactions, *i.e.*, sharing, copyright transfer, borrow GLAM assets etc. to be recorded shall be part of a standard blockchain network (JPEG-Blockchain in this instance). A *transaction hash* will then be incorporated within existing standard metadata of JPEG files. Within the blockchain the transaction blocks will have information related to copyright, ownership, transaction details and a signature/hash of the image. We advocate the content signature is important to verify media integrity and authenticity. Finally a copy of the image with updated metadata is stored in blockchain complaint media server and distributed to the client when a genuine transaction is validated. Detailed description of the proposed architecture is discussed in Section 3. Main contributions of this work are:

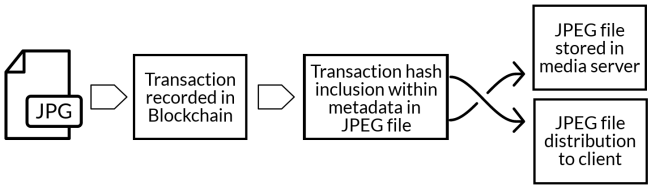

Fig. 1. Flow diagram of the proposed JPEG-blockchain.

- A JPEG-blockchain framework for GLAM services ensuring image integrity and authenticity.
- A proposition for standardization complementing JPEG Privacy and Security framework.
- Content authentication using compressive sensing and highly compressed low resolution image coding.

2 Background

2.1 JPEG Privacy and Security

The JPEG committee has initiated a standardization process to enable privacy and security support in its various standards. This activity is formally known as JPEG Systems Part 4 'Privacy, Security and IPR features' or ISO/IEC 19566-4. JPEG Privacy and Security intends to provide a degree of

trust while sharing image content and metadata. Simultaneously, it will allow the signalling of the associated policies. It targets technical solutions for resolving privacy and security issues, which are compliant with legacy technology in the domain, *i.e.*, both image coding as well as metadata standards that signal information such as access policies and IPR conditions. The work on JPEG Privacy and Security is currently in progress and is expected to become an international standard in 2019.

2.2 Blockchain

Emerging *blockchain* [4] technology is an open distributed *ledger* (database) that records every transactional details referred as *blocks*. Each record or block is timestamped, linked to a previous block and resilient to modification of the data. Therefore blockchain is considered to be a trusted and secured mechanism for transactions between two or more entities in an efficient, verifiable and permanent way. Increasing interests in this technology were noticed in recent years largely from industries and academia that intend to adopt the concept to provide a secure and publicly verifiable transaction mechanism. For example, Hyperledger is a recent umbrella project of open source blockchain and related tools to support the collaborative development of blockchain-based distributed ledgers[1].

Blockchain technology allows transactions to be verified without using a central organisation to process the transaction [5]. Instead multiple nodes are used to form a consensus on whether a transaction is valid or not. An example of blockchain working principle is shown in Fig. 2 where a payment is sent from A to B while other nodes verify the transaction. In case of a transaction failure or invalidation, the transaction is not acknowledged. Eventually all nodes will verify and add the transaction to their copy of the ledger. Conceptually it works by connecting or chaining blocks of information about the transactions and storing them together in a chronological order and hence called *blockchain*.

Beyond digital currency, this technology has major potential usage in maintaining a record of any digital content. Current other potential application scenarios include multimedia transactions [2], hardware and software wallets, compliance and identity and a number of other financial and transaction management applications, such as *smart contracts* [6]. Essentially blockchain is relevant to anything that requires transaction verification or a signature [4] leading to authenticity and trust. However, no major effort was noticed in multimedia applications except a basic blockchain based digital rights management concept introduced by Fujimura *et al.* [7] where the right information was added as part of blockchain transaction. Initial idea of a multimedia blockchain framework was recently proposed in [2]. Advancing the previous work, in this paper we propose a new framework within the scope of

[1] https://www.hyperledger.org/

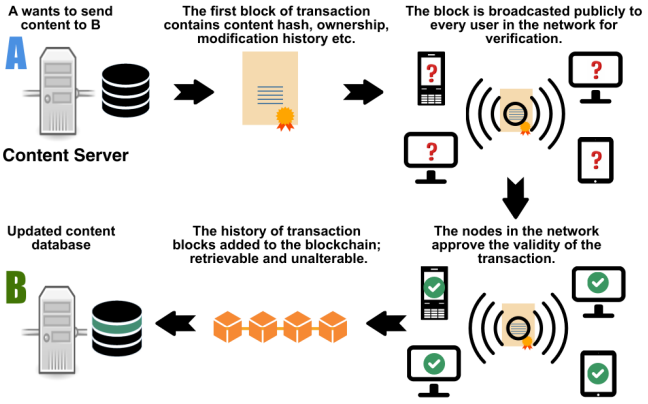

Fig. 2. Overview of the blockchain working principle.

JPEG standard that keeps all records of the media transactions (*e.g.*, ownership, licenses etc.) as well as offers mechanism for tamper-proof verifiable integrity of the media.

2.3 GLAM services

Novel interaction technologies such as tables, walls, combined with personal devices, enable many possibilities for personalized information delivery. The GLAM sector needs to deliver information in a personalized experience, *i.e.*, both online and offline, one where the museum learns about the customer in order to present them with relevant information on each visit. This has the potential to increase return visits by tailoring the information presented in exhibits and expanding on this, or giving a different perspective, in future visits. There are many emerging digital technologies that can be used for this, although may privacy and user interaction challenges arise.

The GLAM sector is dealing with images that have associated Intellectual Property Rights (IPR). In this particular case, the access to the images needs to be controlled based on specific privacy policies or rules and IPR conditions. Therefore, mechanisms are needed to both specify the policies and enforce them. As an example, implementation of controlled access to images. One relevant issue here is who defines such policies. Two main options are either the museum or the image owner, assuming the case in which the image owner is not the one that provides access to the image. In our example, the the museum could define the rules, but they should be based on the original intellectual property rights set by image owner. Such issues are well handled Within the Blockchain framework in a manner by which the image owner can create a smart contract with the museum to provide relevant access control to these images.

The next issue is how to define the rules (policies); *i.e.*, on which kind of information we should base the conditions to verify in order to decide if the access is granted. Examples could include information on user, image, action, context, etc.

A specific example of a rule to illustrate this could be only museum employees can view the painting photo album during this month. In this case, users are the museum employees, images are the specified photo album, action is view and context is this month.

The third and final issue we consider is where to store the rules and the images to which they apply. Images could be even referenced in their repository. Rules could be included as protected metadata in the image itself using JPEG Privacy and Security signaling syntax or in external systems. Blockchain provides a sophisticated and advance approach to rule based content management.

3 The JPEG-blockchain framework

3.1 Expected features

JPEG Privacy and Security is a new standard which is defined for the family of JPEG standards to increase the reliability of encoded images and associated metadata. Expected features of the standard enable applications and services that have better protection ability of image content on digital publishing, image sharing, and content distribution via the Internet. Those features are basically classified into two general categories: protection and authenticity. An important protection feature is to utilize protection tools such as encryption and watermarking technologies to protect parts of any type of JPEG images and/or associated metadata independently. Regarding the authenticity feature, it is an essential feature in many use cases to ensure and check the integrity of image data and/or embedded metadata. Together these features also enables trust among the users in the backdrop of emerging fake image/news related issues.

For the realisation of those expected features, the emerging blockchain technology could provide a tractable and scalable solution through a distributed and tamper-proof media transaction framework. In this work we propose a JPEG-blockchain framework that compliments the upcoming JPEG Privacy and Security standard by providing a decentralised architecture incorporating protection and authenticity, two main features of the standard. In the following subsections we describe the architecture of this framework and how does that fit within the existing narratives of JPEG Privacy and Security.

3.2 JPEG-Blockchain architecture

The proposed distributed and tamper proof JPEG-blockchain framework consists of three parts: *a)* a blockchain network, *b)* blockchain transaction consists of image information including copyright and content signature/descriptor and *c)* a proposition for standardization in JPEG privacy and security standard through JPEG box file format incorporating the blockchain *transaction hash*. A three-node example blockchain network is depicted in Fig. 3 showing media filestores connected to individual nodes to store images. The proposed blockchain transaction information fields are shown in Fig. 4. Finally the linkage between blockchain and the existing JPEG format (using box structure) is shown in Fig. 6 and discussed in Section 3.3.

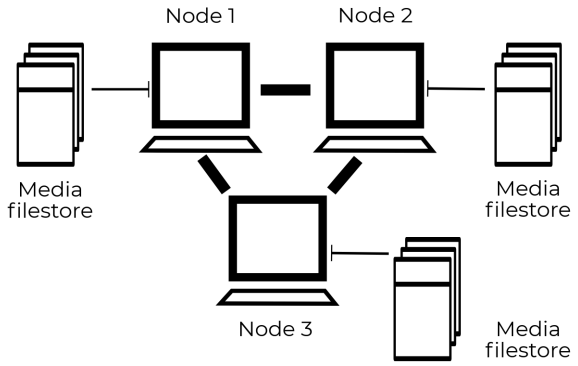

Fig. 3. Overview of the JPEG-blockchain network with three nodes. Media filestores are attached to individual nodes.

3.2.1 The blockchain network

The proposed framework uses a standard blockchain infrastructure and was amended to satisfy the requirements of the proposed framework. As a proof of concept realisation, we implemented a private three-node blockchain network as shown in Fig. 3. *Geth*, a tool provided by Ethereum[2] is used to build this personal network. Essentially the network consists of a number of Ethereum Virtual Machines (EVM) or *nodes* connected to every other node to create a mesh through a standard access point/router. Each EVMs runs an individual copy of the entire blockchain and can performs various operations such as mining a block or validate a transaction. When a new block is added to one of the nodes, the blockchain updates and each node is synchronised.

In the proposed architecture, media file stores are attached to each node. These are used to store a copy of the media that are registered with the blockchain. The filestores are created using a relational database which contains a table with a number of fields related to image information including the blockchain *transaction hash*. The interface allows to link the image related to a specific transaction. It is to be noted that the images are not synchronised across the nodes as that is unnecessary and consumes considerable bandwidth. Rather we rely on the transaction record (discussed in Section 3.2.2) that has sufficient information about the media asset.

3.2.2 The blockchain transaction

The second part of our architecture is the transaction block which carries all necessary information related to the stake-

[2] https://www.ethereum.org/

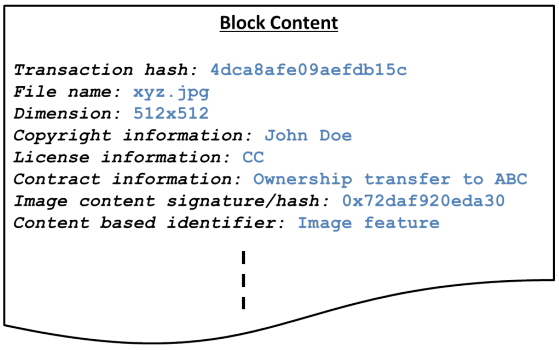

Fig. 4. Proposed information content within a blockchain block for each transaction.

holders, copyrights and the image content. The proposed structure of each blockchain transaction is depicted in Fig. 4. We advocate the fields should include following components,

- A unique *transaction hash* for every media transaction recorded in the blockchain. The transaction type includes copyright transfer, IPR and access control information for the assets owned by museums etc. This unique *transaction hash* is also the link between the blocks that are derived from the current one.
- Basic *asset information* including original file name, dimensions etc.
- *Copyright information*, *e.g.*, ownership, digital rights and licensing,
- *Contract information* relating to ownership transfer, access control information/rules set by the museums or other entities within the GLAM sector,
- *Content signature* that ensures authenticity and integrity of the asset and
- *Content based identifier* consisting visual descriptors which can be captured by the local and global features of the image.

The image signature can be chosen in one of the following ways: *1)* compressive sensed (CS) samples and *2)* a low resolution low quality low data rate bit stream extracted from a JPEG 2000 bitsreatm. We advocate that using the content signature it is plausible to verify the integrity and the authenticity of the content by reconstructing an approximation of the original image [2]. For example, the blockchain contains the content hash which is used to reconstruct an approximate version of the original image. When a test image is received, the system will first generate a similar content hash followed by a reconstruction. A comparison between the original and test approximation information detects any tampering attempt and

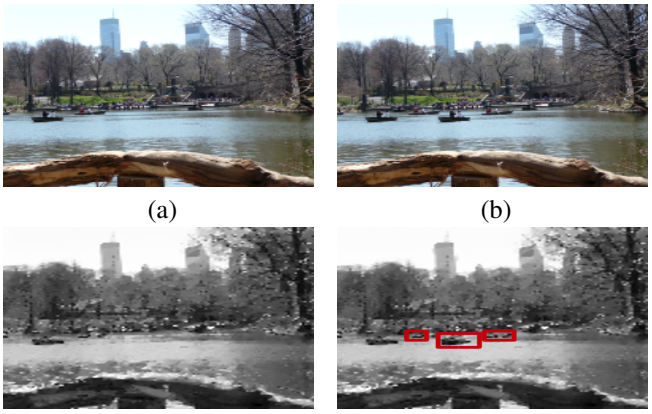

Fig. 5. (a) Original image (b) tampered image (c) reconstructed image using original CS samples and (d) tampered region detection.

hence verify the integrity and authenticity of the received asset. An example of tamper detection using compressive sensing based content hash is shown in Fig. 5. We briefly describe the proposed content hashing techniques within the scope of this framework.

CS samples: The compressive sensing theory proved that it is possible to reconstruct a signal with sparse representation from a reduced set of linear measurements compared to the minimum sampling-rate of Shannon-Nyquist theorem. The standard CS model for a given signal $x \in \mathbb{R}^n$ in the sparse domain can be described as:

$$y = \mathbf{\Phi} x \tag{1}$$

where $\mathbf{\Phi}$ is the sensing matrix with $m \times n, m \ll n \,\&\, y \in \mathbb{R}^m$.

The sparse signal is consisted of small number of non-zero coefficients. Hence, the dense image I usually required a sparsifying transform, *e.g.*, DFT, DCT or DWT with basis functions $\mathbf{\Psi}$ to achieve a more compact energy distribution of the signal [8].

$$x = \mathbf{\Psi} I \tag{2}$$

The reconstruction is usually a non-linear operation to reconstruct an approximation of the original signal. Since the optimization constraints of reconstructions are different, some algorithms can reconstruct the image domain values without sparsifying the transform in Eq. (2). The approximate reconstruction (lossy) in this work is sufficient to detect tampering.

Highly compressed low resolution low bit rate image stream: Alternative to the compressed sensed samples, we also propose to use low resolution, low bit rate image stream derived from JPEG 2000 coding standards [9] or similar where one can downsample the image to a lower resolution

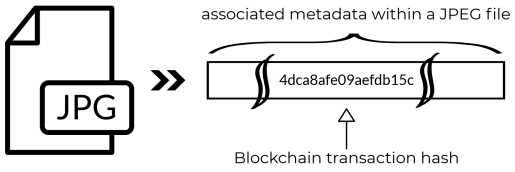

Fig. 6. Proposed blockchain *transaction hash* information inclusion within the metadata contained in a JPEG file.

Field name	LBox	TBox	XLBox	DBox
Size(bits)	32	32	0 or 64	Variable
Meaning	Box Length	Box Type	Box Extended Length (optional)	Box Contents

Fig. 7. Organization of box structure

and then compress by any image coding algorithm. For example Wu *et al.* [10] proposed a low bit-rate image compression via adaptive down-sampling which can be used as the content hash within the blockchain transaction. A reconstruction can be performed using a constrained least squares upconversion for authentication purposes.

Content based identifier: In addition to having provisions for image integrity verification, we propose that the transaction block includes a field for feature based visual descriptors for content based asset identification. This feature will enable fast visual search without accessing the original image. We suggest using a MPEG-7 like Compact Descriptors for Visual Search (CDVS) [11]. The descriptor can be low level local features such as compressed histogram of gradients [12], grid-based quantization coding [13] etc. or global descriptors, *e.g.*, visual bag-of-words based signature aggregating the statistics of local descriptors [14]. We argue that the descriptor should be limited to a small number of bytes, hence suitable for a blockchain transaction.

3.3 Proposition for standardization in JPEG Privacy and Security

Considering the blockchain *transaction hash* is the primary linkage between the digital asset (JPEG images in this case), we propose to include this information within the metadata encapsulated in a JPEG file as shown in Fig. 6. Our emphasis is on the upcoming JPEG Privacy and Security standard as discussed in Section 2.1. Our proposition considers the box structure in JPEG as discussed below.

Box structure in JPEG A box-based file format defined in ISO/IEC 15444-1 known as JPEG 2000 Part 1 is high flexible and extensible since it provides a simple parsing mechanism for file readers to extract necessary information. The box-based file format is composed of a collection of boxes and each box consists of four elements as depicted in Fig. 7. **LBox** in Fig. 7 specifies the length of the box. **TBox** defines the type of information contained in **DBox** field. The value of **TBox** field is typically described by 4CC defined in ISO/IEC 646, *e.g.*, 'jp2c (0x6A70 3263)'. If the size of a box length exceeds the capacity of 32 bits representation, **XLBox** field can alternatively be used to store the extended length in capacity of 64 bits representation.

The box-based format has been extended by introducing the concepts of both marker segment and box format so that boxes can be integrated into an application marker defined in ISO/IEC 10918-1. More precisely, $\mathbf{APP}_{11}$ marker is reserved as JPEG XT Marker in ISO/IEC 18477-1, which is a subset of ISO/IEC 10918-1. The binary structure of the $\mathbf{APP}_{11}$ marker segment for boxes are depicted in Fig. 8.

Since the size of marker segment is limited to 64KB, box contents would be divided into multiple fragments. If box contents are separated, **Payload Data** stores each of the fragments and **Z** counts the fragment number. File readers shall collect all the associated marker segments to concatenate the box contents. More detailed definitions are described in ISO/IEC 18477-3.

Based on the flexible binary structure and backward compatibility of the box based file format, we propose that a unique *transaction hash* is encapsulated within a box in our proposed blockchain framework. The box can be defined by combination of a new box type and a variable parameter for the *transaction hash*. For instance, the box type of the *transaction hash* could be 'BCTH (0x4243 5448)' and the associated parameter H could specify the *transaction hash* as a variable binary field.

3.4 The JPEG-blockchain for GLAM

Our idea is to have a blockchain network with multiple nodes associated with the various galleries, libraries, archives and museums (termed as *entity*) across a region, country, continent or universe. When a specific entity initiates a transaction, *i.e.*, passing digital asset to another entity, requires to record and verify the transaction in the blockchain. That implies the network *JPEG-blockchain* will record all transactions with relevant information including copyright, access control, content signature (for authenticity), visual descriptor etc. and synchronise with all other nodes in order to validate the transaction. Once validated the transaction will provide a unique *transaction hash* which needs to be incorporated within the metadata in the JPEG file following the proposed compliance in the standard. This modified JPEG file/asset, is then updated in the media filestore and distributed between two entities. The recipient can also verify the integrity of the asset through a) validation of the blockchain record and b) comparing content signatures between record in the blockchain and the signature derived from the received asset.

Field name	APP11 (0xFFEB)	Le	CI (0x4A50)	En	Z	LBox	TBox	XLBox	Payload Data
Size(bits)	16	16	16	16	32	32	32	0 or 64	Variable
Meaning	Identifies all JPEG XT Marker Segments	Length of the marker segment	Common Identifier	Box Instance Number	Packet Sequence Number	Box Length	Box Type	Box Extended Length (optional)	The syntax of the concatenated payload data is defined in each JPEG standard

Fig. 8. Binary structure of JPEG XT marker segment

While this proposed architecture is aimed at GLAM sector due to its emerging need as outlined in the beginning of this paper, the architecture is flexible and scalable to accommodate any other sectors, *e.g.*, stock image market, social media sharing / social network services, etc. We envisage a single distributed JPEG-blockchain network with multiple nodes across geographies. Due to its decentralised architecture (Section 2.2), no single entity or organisation can claim ownership of the network and hence offers a transparent and trusted single media transaction framework. The proposition for standardization within JPEG will provide a platform empowering interoperability and compliance.

4 Conclusions

In this paper, we propose a new blockchain architecture for seamless and trusted media distribution in GLAM sector. The distributed architecture of this framework offers solution to GLAM sector issues related to digital rights management, access control and verification of integrity and authenticity of the media. We also proposed the linkage between the new blockchain architecture and emerging JPEG Privacy and Security standard.

We demonstrate that while maintaining backward and forward compatibility with legacy image coding standards such as JPEG, the blockchain *transaction hash* can easily be integrated within the family of coding technologies standardized by the JPEG committee. The proposed solution is in the process of integration in the JPEG Systems Part 4 'Privacy, Security and IPR features (ISO/IEC 19566-4) and Part 5 'JPEG Universal Metadata Box Format (JUMBF) (ISO/IEC 19566-5) standards. We illustrated, how the use of blockchain technology can be used for distributing digital media content effectively within the GLAM supply chain.

5 References

[1] F. Temmermans, T. Ebrahimi, S. Foessel, J. Delgado, T. Ishikawa, A. Natu, and P. Schelkens, "JPEG privacy and security framework for social networking and glam services," *EURASIP Journal on Image and Video Processing*, vol. 2017, no. 1, p. 68, 2017.

[2] D. Bhowmik and T. Feng, "The multimedia blockchain: A distributed and tamper-proof media transaction framework," in *Proc. IEEE Int'l conf. on Digital Signal Processing*, 2017, pp. 1–5.

[3] ISO/IEC JTC 1/SC29/WG 1. Privacy and security final call for proposals. Available: https://jpeg.org/items/20170403_cfp_privacy_security.html [Accessed: April 3, 2017].

[4] W. Mougayar, *The Business Blockchain: Promise, Practice, and Application of the Next Internet Technology*. John Wiley & Sons, 2016.

[5] J. Yli-Huumo, D. Ko, S. Choi, S. Park, and K. Smolander, "Where is current research on blockchain technology? A systematic review," *PloS one*, vol. 11, no. 10, p. e0163477, 2016.

[6] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," in *Security and Privacy (SP), 2016 IEEE Symposium on*. IEEE, 2016, pp. 839–858.

[7] S. Fujimura, H. Watanabe, A. Nakadaira, T. Yamada, A. Akutsu, and J. J. Kishigami, "BRIGHT: A concept for a decentralized rights management system based on blockchain," in *IEEE 5th International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, 2015, pp. 345–346.

[8] J. Romberg, "Imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 14–20, 2008.

[9] M. W. Marcellin, M. J. Gormish, A. Bilgin, and M. P. Boliek, "An overview of JPEG-2000," in *Proc. Data Compression Conference (DCC)*, 2000, pp. 523–541.

[10] X. Wu, X. Zhang, and X. Wang, "Low bit-rate image compression via adaptive down-sampling and constrained least squares upconversion," *IEEE Trans. Image Processing*, vol. 18, no. 3, pp. 552–561, 2009.

[11] L.-Y. Duan, J. Lin, J. Chen, T. Huang, and W. Gao, "Compact descriptors for visual search," *IEEE MultiMedia*, vol. 21, no. 3, pp. 30–40, 2014.

[12] V. Chandrasekhar, G. Takacs, D. Chen, S. Tsai, R. Grzeszczuk, and B. Girod, "Chog: Compressed histogram of gradients a low bit-rate feature descriptor," in *Proc. IEEE CVPR*, 2009, pp. 2504–2511.

[13] B. Girod, V. Chandrasekhar, D. M. Chen, N.-M. Cheung, R. Grzeszczuk, Y. Reznik, G. Takacs, S. S. Tsai, and R. Vedantham, "Mobile visual search," *IEEE signal processing magazine*, vol. 28, no. 4, pp. 61–76, 2011.

[14] R. Ji, L.-Y. Duan, J. Chen, H. Yao, J. Yuan, Y. Rui, and W. Gao, "Location discriminative vocabulary coding for mobile landmark search," *International Journal of Computer Vision*, vol. 96, no. 3, pp. 290–314, 2012.