Distributed Video Coding: Selecting the most promising application scenarios

Fernando Pereira [a,*], Luis Torres [b], Christine Guillemot [c], Touradj Ebrahimi [d], Riccardo Leonardi [e], Sven Klomp [f]

[a] *Instituto Superior Técnico– Instituto de Telecomunicações, Portugal*
[b] *Technical University of Catalonia, Spain*
[c] *INRIA, France*
[d] *Ecole Politechnique Fédérale de Lausanne, Switzerland*
[e] *Università degli Studi di Brescia, Italy*
[f] *Leibniz Universität Hannover, Germany*

ARTICLE INFO

ABSTRACT

Distributed Video Coding (DVC) is a new video coding paradigm based on two major Information Theory results: the Slepian–Wolf and Wyner–Ziv theorems. Recently, practical DVC solutions have been proposed with promising results; however, there is still a need to study in a more systematic way the set of application scenarios for which DVC may bring major advantages. This paper intends to contribute for the identification of the most DVC friendly application scenarios, highlighting the expected benefits and drawbacks for each studied scenario. This selection is based on a proposed methodology which involves the characterization and clustering of the applications according to their most relevant characteristics, and their matching with the main potential DVC benefits.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Distributed Video Coding (DVC) is a new video coding paradigm based on two major Information Theory results which set the foundations of Distributed Source Coding (DSC): the Slepian–Wolf [21] and Wyner–Ziv theorems [23,24]. The Slepian–Wolf theorem [21] addresses the case where two statistically dependent, discrete random sequences, X and Y in Fig. 1, are independently encoded, and not jointly encoded as in the largely deployed predictive coding solution adopted in MPEG and ITU-T video coding standards. 'Surprisingly', the theorem says that the minimum rate to encode the two dependent sources is the same as the minimum rate for joint encoding, with an arbitrarily small probability of error, when joint decoding is performed and the two sources have certain statistical characteristics, notably are jointly Gaussian. Later, it would be shown that only the innovation, this means the X–Y difference, needs to be Gaussian, relaxing the requirements on the joint X and Y statistics. This is an important result because it opens the doors to a new coding paradigm where, at least in theory, separate encoding does not induce any compression efficiency loss when compared to the joint encoding used in the traditional predictive coding paradigm (see both paradigms in Fig. 1).

Slepian–Wolf coding is the term generally used to characterize coding architectures that follow the scenario described in Fig. 1b. Slepian–Wolf coding is also referred to in the literature as lossless distributed source coding since it considers that the two statistically dependent sequences are perfectly reconstructed at a joint decoder (neglecting the arbitrarily small probability of decoding error), thus approaching the lossless case. Slepian–Wolf

---
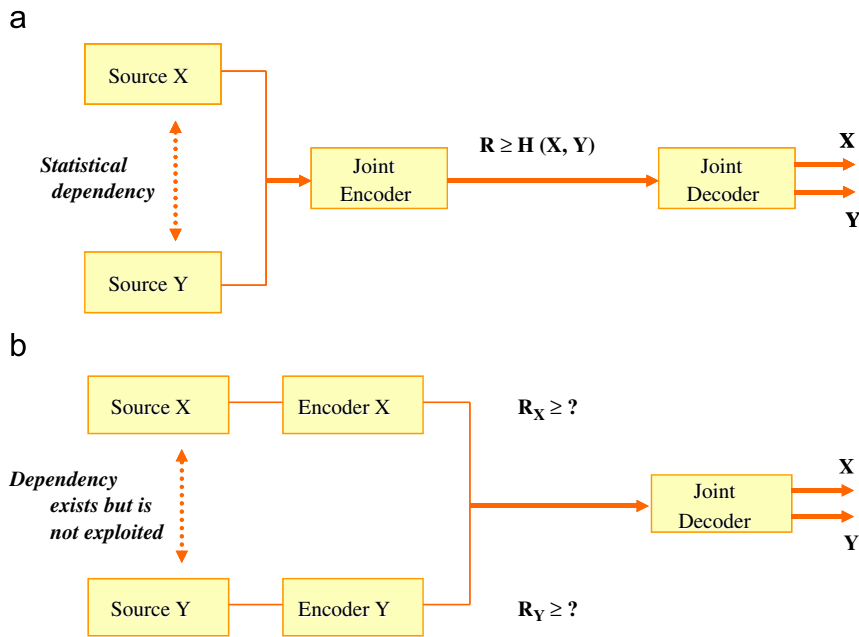
Fig. 1. Conventional predictive versus distributed video coding paradigms: (a) joint encoding and joint decoding as in current coding standards and (b) independent (distributed) encoding and joint decoding.

coding has an interesting relationship with channel coding: the dependency between the X and Y sequences can be modeled as a virtual 'dependency' channel where X represents the original uncorrupted information, while Y is used to estimate a noisy version of X designated as side information. The 'estimation errors' between X and the side information obtained from Y (sometimes designated as Y itself) can be corrected by applying channel coding techniques to encode the X sequence where Y plays the role of systematic information in regular channel coding. There is thus no surprise that channel coding tools typically play a main role in the new video coding paradigm.

However, there is a major constraint in the Slepian–Wolf theorem since it refers to lossless coding and this is not the 'most exciting' case in practical video coding solutions; this comes from the fact that lossless coding achieves rather small compression factors since it does not eliminate the irrelevant video information unperceivable for the human visual system. Fortunately, in 1976, Wyner and Ziv [24] studied the corresponding lossy coding case and derived the so-called Wyner–Ziv theorem. This theorem states that when performing independent encoding there is no coding efficiency loss with respect to the case where joint encoding is performed, under certain conditions, even if the coding process is lossy (and not lossless anymore).

Together, the Slepian–Wolf and the Wyner–Ziv theorems suggest that it is possible to compress two statistically dependent signals in a distributed way (separate encoding, joint decoding) approaching the coding efficiency of conventional predictive coding schemes (joint encoding and decoding). When applied to video coding, this new coding paradigm—Distributed

Source Coding—is well known as Distributed Video Coding (DVC) or Wyner–Ziv (WZ) video coding and opens the doors to new video coding architectures addressing new application requirements. Although the theoretical foundations of DVC have been established in the 1970s, the design of practical DVC schemes has been proposed only in recent years [1,4,18]. A major reason behind these latest developments is related to the evolution of channel coding, notably the emergence of turbo and Low-Density Parity-Check (LDPC) coding, which provides ways to build the efficient channel codes necessary for DVC.

The analysis of the DVC basics, its associated statistical approach, and the main DVC practical approaches available allow concluding that DVC-based architectures may present the following functional benefits [1,4,18,21,23,24]:

1. *Flexible allocation of the overall video codec complexity.* Since the DVC approach allows moving part of the encoder complexity to the decoder, it may provide the benefit of a flexible allocation of the video codec complexity between the encoder and decoder. This means that a codec may decide to dynamically 'invest' some or less complexity at the encoder or the decoder, e.g. for correlation noise modeling or motion estimation, depending on the particular conditions at hand, e.g. the currently available battery at both sides or the available bandwidth. A particular case of this flexible allocation is the important case of low encoder complexity which may also imply lower encoder battery consumption, as well as cheaper and lighter encoders. It is important to stress that 'low encoder complexity' is a moving target and a 'slippery road' since what is complex today may not be complex

anymore tomorrow. However, from a relative perspective, what is less complex today should still be less complex tomorrow and, for example, new devices may be able to accommodate better the 'lower complexity' solution.

2. *Improved error resilience.* Since DVC codecs do not rely on the usual encoder prediction loop but rather on a statistical approach, the propagation in time of errors due to channel corruption is less critical; in fact, DVC behaves as a joint source-channel coding solution where the bits spent work simultaneously to improve quality and recover from errors.

3. *Codec independent scalability.* While in current scalable codecs there is typically a predictive approach from lower layers to upper layers, requiring the encoder to always know what are the decoding results for previous layers in order to create the successive enhancements, the DVC prediction loop free approach between the scalable layers does not require a deterministic knowledge on the previous layers (just a correlation model) which means the layers may be generated by various, different and unknown codecs. This codec independence may also be extended in terms of spatial resolution, which means there is freedom to enhance a video sequence to a certain spatial resolution, starting from any lower (and even unknown) spatial resolution.

4. *Exploitation of multiview correlation.* While the benefits listed above apply for monoview video sequences, there are also functional benefits when a DVC approach is used in a multiview video context where there exists inter-view correlation between the various views of the same scene. In this case, the DVC approach provides a significant architectural benefit since while a typical predictive approach will exploit the inter-view correlation at the joint encoder requiring the various views to be simultaneously available at some encoder location, and thus the various cameras to communicate among them, DVC-based encoders do not need to jointly process the various views and thus do not need inter-camera and inter-encoder communication. It is also important to stress that, in this case, the alternative standards based coding solution implies the independent coding of the various cameras which makes easier for a DVC-based solution 'to beat' it from a compression efficiency point of view.

It is nowadays more and more accepted that the DSC principles are leading to varied tools which may help to solve different problems, e.g. coding [1,4,18], authentication [14], and secure biometrics [3]. While it is difficult to state, at this stage, if any video coding product will ever use DSC principles and how, it is most interesting to study and research this possibility; this is the main target of this paper which adopts a functional point of view for its study.

The functional benefits listed above (called in the following *main DVC benefits*) will be helpful in this paper to select the most promising DVC applications using the methodology presented in the next section. Many of the

DVC advantages discussed along this paper are valid under the assumption that research will bring some major performance developments in the next years; e.g. in terms of compression efficiency this is already happening. Although the literature generally refers that DVC is mainly useful for low complexity and low-power consumption encoders, no detailed application analysis is available on these benefits [4]. It is also believed by the authors that low complexity is not the single potential DVC benefit, and may not even be the most promising one (see comments above). This investigation is precisely one of the major current DVC research targets.

In the literature, several application scenarios are typically presented as those mainly benefiting from the new coding approach but no exhaustive study has been performed [4]. The major objective of this paper is to study in a more systematic way which are the application scenarios for which the DVC paradigm may bring major benefits and identify what are these benefits. Note that it is not the purpose of this paper to claim that DVC is the right way to go for any application. Considering the far from mature stage of DVC research, it is too early for final conclusions and claims. The purpose is rather to identify the most promising application scenarios, helping the researchers to focus their work on the most adequate application spots, in order that conclusions on the real value of DVC for these applications may be reached as soon as possible. To achieve the objective stated above, this paper proposes a selection methodology which involves listing the DVC potential advantages, for example, error resilience, flexible encoder–decoder complexity trade-offs, and multiview video, as well as the current DVC drawbacks, for example, coding efficiency[1] and decoder complexity. Afterwards, the application scenarios are clustered according to various relevant characteristics, e.g. single/multiple cameras, availability of a return channel, encoder/decoder critical complexity, delay constraints, and, finally, a list with the application scenarios for which DVC looks to be more promising will be drawn, based on the proposed methodology.

In summary, Section 2 proposes a methodology to select the most promising DVC application scenarios and lists the application scenarios selected for analysis, while Section 3 identifies the DVC potential advantages and current drawbacks for each applications scenario. Section 4 characterizes and clusters the applications scenarios and, finally, Section 5 selects the most promising DVC application scenarios. Section 6 concludes the paper.

2. Methodology towards the most promising DVC application scenarios

For the purpose of this paper, a significant number of scenarios have been considered, avoiding the up front elimination of any interesting scenario. Naturally, some of

---

[1] Although theoretically DVC may reach the same compression efficiency as predictive coding (under certain conditions), practical solutions are not yet mature enough to reach this level of compression efficiency; however, there are already interesting solutions if a compression efficiency–complexity trade-off is considered [2,17].

them will be more promising than others as it will be concluded at the end of this paper. There may be some overlapping between the various application scenarios addressed since the boundaries between scenarios are sometimes fuzzy. While the consideration of a high number of application scenarios may create some redundancy between them, this rather exhaustive approach was adopted since it guarantees that nothing major should be missed. Having studied a long list of scenarios, this should help guaranteeing that the paper's conclusions are solid and meaningful.

Although there are various approaches to achieve the objectives stated above, this paper proposes the following methodology:

1. *Identification of the list of application scenarios to analyze*, e.g. by reviewing the literature, to gather those which have been considered relevant by DVC researchers; this list must be rather complete and exhaustive (see later in Section 2).
2. *Listing of the DVC potential advantages and current drawbacks for each application scenario* by matching the application requirements with the DVC pros and cons claimed in the literature (see Section 3).
3. *Clustering of the application scenarios based on some major characteristics* (see Section 4).
4. *Selection of the most promising DVC application scenarios* for the various relevant application clusters previously identified (see Section 5):
   a. Counting for each application scenario the number of DVC potential advantages matching the already identified DVC main benefits; it is assumed that DVC research will reduce the impact of the identified drawbacks, e.g. the DVC coding efficiency gap to predictive coding will be shortened.
   b. Selection as most promising application scenarios of those with the highest count in 4.a while maximizing at the same time the coverage of the clusters identified in 3), this means at least one application per cluster will be selected.

Following the proposed methodology, the list of DVC relevant application scenarios selected for detailed analysis in this paper is:

1. wireless video cameras;
2. wireless low-power surveillance;
3. mobile document scanner;
4. video conferencing with mobile devices;
5. mobile video mail;
6. disposable video cameras;
7. visual sensor networks;
8. networked camcorders;
9. distributed video streaming;
10. multiview video entertainment;
11. wireless capsule endoscopy.

This list shows an assortment of applications, with some overlapping among them, notably real-time and non-real-time systems, unidirectional and bidirectional, monoview and multiview, different complexity and battery constraints, etc. The following sections will study in detail each of the selected application scenarios using the methodology described in the previous section.

3. DVC advantages and drawbacks by application scenario

This section performs a detailed analysis of each relevant application scenario in terms of potential DVC advantages and current DVC drawbacks. In practice, the DVC advantages should correspond to potential DVC benefits that most of the times may only become effective if the drawbacks are removed or, at least, significantly reduced. This is clearly the case regarding the coding efficiency gap which has been reduced in recent years in many ways, e.g. by improving the side information creation, and the correlation noise modeling [2,17]. Although a DVC-based system may not need to provide precisely the same rate-distortion (RD) performance as standards based coding systems to be commercially deployed, it must for sure provide a good enough trade-off between advantages and drawbacks regarding alternative solutions.

3.1. Wireless video cameras

An important application scenario for DVC is related to the wireless communication of video signals between remote devices. With the new emerging technologies for wireless communication, the possibility of sending video data in a wireless fashion has become a reality. This section mainly addresses the use of single wireless cameras; the situation where a single wireless camera has to send the acquired video data to a central station is the most relevant. Although an important application for this type of cameras is surveillance, this case will not be considered here since it will have a specific section in the following.

The first example application deals with the possibility of using small portable cameras for video gathering in diverse situations, e.g. meetings, parties, etc. (see Fig. 2, left). Also, this type of cameras can be integrated in embedded systems for cars, trains, airplanes, or any mobile environment. In those situations, the use of a wireless camera is the only viable choice because it is often not possible to use a wired solution, especially if the user wants to have a highly flexible system where the camera can be easily moved from one place to another.

Another interesting application is the case of very small wireless cameras for police investigation purposes or for remote sensing of phenomena that are very hard to be physically reached. In both situations, one needs to send a video signal from one point to a station while using very small devices and thus with very limited resources. Finally, wireless cameras also have great value in television production environments, being much used both inside and outside the studio to avoid annoying cables.

Table 1 presents the most relevant potential DVC advantages and current DVC drawbacks for wireless video

cameras applications. Since this type of table will be presented for each application analyzed, each potential advantage or current drawback will only be detailed the first time it appears; afterwards, only new advantages and drawbacks, or specific relevant comments for each application scenario will be added.

3.2. Wireless low-power surveillance

Wireless low-power surveillance network applications are mainly about surveillance, and therefore security.

Fig. 2. (Left) ordinary wireless camera [10] and (right) wearable wireless webcam imitates surveillance cameras common in casinos and department stores [7].

With this purpose, various low-power consumption components are interconnected and the communication between them is carried out through wireless communication protocols. The components that provide information to the system are cameras (although other sensors can also be present) and the images can be captured or displayed by one or multiple devices. While some wireless surveillance applications consider only a single camera, other applications consider a multiview scenario where there is inter-view correlation to be exploited. Also, quality and spatial scalability may be relevant issues if the decoder for a specific view uses the image decoded from another view to provide decoded video with increased quality or resolution. In this case, the additional quality or resolution will be provided based on data which is not deterministically known.

Among some of most important low-power surveillance applications are traffic monitoring, surveillance inside transports and taxis [12], electronic tagging (a form of non-surreptitious surveillance consisting of an electronic device attached to a person or vehicle allowing their whereabouts to be monitored), wireless home monitoring, wildlife and fire monitoring, military reconnaissance and monitoring, sousveillance (refers to the recording or monitoring of real or apparent authority figures by others, particularly those who are generally the subject of surveillance; see Fig. 2, right) [9]. Table 2

Table 1
DVC potential advantages and current drawbacks for wireless video cameras

DVC potential advantages	DVC current drawbacks
Lower encoding complexity DVC has received a lot of attention in recent years because it offers the possibility of shifting computational complexity from the encoder to the decoder. There are already available DVC codecs which provide interesting rate-distortion (RD) performance–encoder complexity trade-offs, notably regarding H.264/AVC Intra coding.	*Higher decoding complexity* One of the main DVC characteristics is the potential to shift the complexity from the encoder to the decoder. In current DVC approaches, the required decoding complexity seems to be rather high; in applications requiring real-time decoding, this may be a significant drawback (that should become less relevant with time).
Lower size and weight devices As the complexity of the encoder is supposed to be reduced with the DVC approach, the size and the weight of the devices that capture the video may also be reduced; this is relevant for the type of application scenarios addressed in this section.	*Lower compression efficiency* Until now, DVC did not reach the same level of compression efficiency as state-of-the-art predictive coding, notably the H.264/AVC standard. However, for lower complexity encoding there are already interesting solutions, e.g. providing a RD performance better then H.264/AVC Intra or even H.264/AVC zero motion with lower complexity [2,17]. Since there is a growing interest in DVC research, it is also expected that the DVC RD performance will improve substantially, thus eliminating or at least significantly reducing this drawback.
Lower encoding power consumption In addition, the lower encoder complexity may reduce the power consumption, which means longer battery life or reduced battery size, or more power available for transmission and thus higher transmission range.[a]	
Improved error resilience It is well known that the predictive video coding approach is strongly affected by channel errors propagation. It has already been shown that a DVC approach may be more suitable, as no prediction loop is used and thus no prediction error is sent [16,20]. The prediction in the standard encoding phase is substituted by the side information inter(extra)polation at the decoder in the distributed approach; as long as the decoder has good side information, the original signal is recovered regardless of the presence of previous errors, provided enough WZ bits are received from the encoder.	

[a] It is worthwhile to note that although the 'lower encoding complexity', 'lower size and weight devices' and 'lower encoding power consumption' advantages are closely related, it is meaningful to explicitly mention them since there are application scenarios where the three advantages are not equivalent; for example, there are applications where low encoding complexity is a need but low size and weight are not relevant.

Table 2
DVC potential advantages and current drawbacks for wireless low-power surveillance

DVC potential advantages	DVC current drawbacks
Lower encoding complexity	*Lower compression efficiency*
	The flexible allocation of codec complexity may decrease the impact of this drawback if more complexity may be allocated to the encoder and thus better RD performance is achieved. The typical high temporal correlation of surveillance content, e.g. video from static cameras, may build an easier case for DVC to reduce faster the compression efficiency gap with conventional coding solutions.
Lower size and weight devices	
Lower encoding power consumption As the amount of energy can be limited in some surveillance scenarios, lower consumption impacts on many aspects, from the amount of information to process at the encoder to the volume of wireless communication that can be carried across long distances. As devices' life is longer and less energy and maintenance are required, it allows monitoring harder to reach areas.	*Need for a (network) transcoder* In an end-to-end wireless low-power surveillance network scenario, a transcoder inside the network must be used in order to keep both the encoder and the decoder as simple as possible (the transcoder has to encode the video with a conventional video codec).
Flexible allocation of codec complexity DVC capability of balancing the complexity between encoder and decoder provides flexible solutions to the many different applications that fall in this scenario. This allocation of complexity may be dynamic in time, e.g. may be made dependent on the available battery.	
Improved error resilience This advantage may even be more evident for this application scenario due to the typical high temporal correlation present in video sequences captured with static cameras.	
Multiview correlation exploitation In the case multiple cameras cover the same scene, DVC may exploit the inter-view correlation, notably without requiring the various cameras to communicate among them but rather only with the central control (decoding) node.	
Codec/resolution independent scalability Finally, if some scalability is to be provided based on the inter-view correlation, DVC allows performing this without knowing precisely the data decoded in the lower layers as required for conventional scalable solutions.	

presents the most relevant potential DVC advantages and current DVC drawbacks for wireless low-power surveillance.

3.3. Mobile document scanner

The advent of wireless networks and mobile computing has freed businessmen from their offices, allowing them to work on the go. However, some services remain only available at fixed locations. Among them are copy machines, fax machines, and image scanners. The large volumes and heavy weights of these machines prevent them from being carried along. This issue needs to be alleviated to allow a truly 'anywhere, anytime' working environment. One solution would be to enable mobile phones to be used as portable faxes or scanners that can be used any time, simply by sweeping the phone across the page.

Document scanning on the go with a mobile phone would give wireless carriers the opportunity to provide a host of new services, ranging from the most basic ones like document transmission to email addresses, to printers or to the user's computer, to more advanced services like Optical Character Recognition (OCR) and instantaneous translation for global travelers, sending back the translated text via instant messaging (see Fig. 3). It would also allow queries into remote databases, a service most useful to law-enforcement units trying to collect evidence and identify criminals on the spot.

Scanning an A4-sized page by moving a mobile phone video camera over the document is likely to take about 3–5 s. Assuming a video frame rate ranging from 5 to 10 frames per second, this is going to produce between 15 and 50 images which a central server must merge together in order to extract the text and record any images. The application run on the central server must then forward the processed document to the targeted end device, e.g. e-mail, user's computer, printer, and mobile phone.

Table 3 presents the most relevant potential DVC advantages and current DVC drawbacks for the mobile document scanner application.

3.4. Video conferencing with mobile devices

Video conferencing mostly regards the transmission of synchronized image (video) and speech (audio) back and forth between two or more physically separate locations
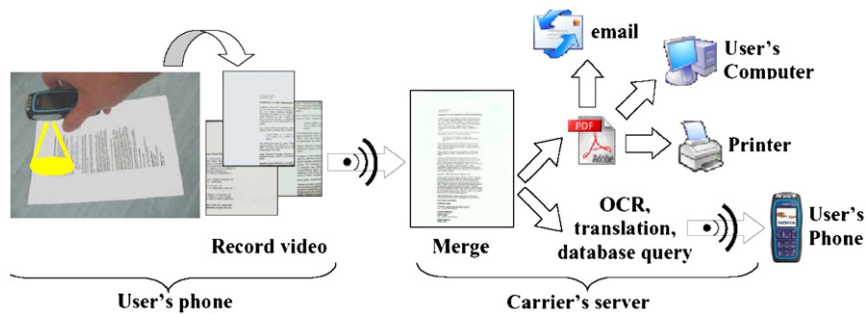
Fig. 3. Document scanning on the go [15].

Table 3
DVC potential advantages and current drawbacks for mobile document scanner

DVC potential advantages	DVC current drawbacks
Lower encoding complexity To reduce complexity, one could consider intra coding (e.g. JPEG or JPEG2000) with a reduced frame rate. However, if the frame rate is too low, this is likely to have an impact on the quality of the reconstructed document. DVC would allow increasing the frame rate and sending extra data. *Improved error resilience*	*Lower compression efficiency* For this type of applications (and content), DVC compression efficiency starts to be, at least, as efficient as H.264/AVC Intra while asking for lower complexity [17]. *Higher decoding complexity* In such applications, since the decoding is performed at a central server, one can afford to have an increased decoder complexity, up to a point related to the scalability of the service, or its capability to support a certain number of users. However, approaches with a more flexible load balancing between encoder and decoder might be beneficial for such applications.

Fig. 4. (Left) video conferencing screen [13] and (right) CVS disposable video camera [8].

(see Fig. 4, left). Sometimes, it is just not possible or practical to have a face-to-face meeting with two or more people. At other times, a telephone conversation or conference call is adequate. Video conferencing adds another possible alternative. Video conferencing should be considered when: (i) a live conversation is needed; (ii) visual information is an important component of the conversation; (iii) the parties of the conversation cannot physically come to the same location; and (iv) the expense or time of travel is an issue.

Table 4 presents the most relevant potential DVC advantages and current DVC drawbacks for video conferencing with mobile devices.

3.5. Mobile video mail

The interest of customers for the new features of mobile devices is growing continuously. Recent statistics show that every 2 years a mobile phone is replaced by a more modern one, allowing new applications, which had not been supported before. One of the most popular applications is sending text messages to friends, family, or fellow-workers if direct calls are not possible or desired.

The first such application was the Short Message Service (SMS), a service for transmitting text messages developed for GSM mobile networks. As a successor of SMS, Multimedia Messaging Service (MMS) was

Table 4
DVC potential advantages and current drawbacks for video conferencing with mobile devices

DVC potential advantages	DVC current drawbacks
Lower encoding complexity DVC lower encoding complexity may make smaller and cheaper devices possible.	*Lower compression efficiency* The fact that videoconferencing video shows a high temporal correlation eases the reduction of the DVC efficiency gap regarding conventional video coding.
Lower encoding power consumption *Increased resolution for same complexity* Alternatively to lower complexity, the resolution of the captured video may be increased while power consumption or computational complexity is maintained. *Improved error resilience*	*Need for a (network) transcoder* In this scenario, a transcoder in the network has to be used in order to also keep the decoder as simple as possible (the transcoder has to encode the video with a conventional video codec). This might be a bottleneck in future developments, namely in terms of total end-to-end delay. A relevant research challenge may thus be the development of efficient DVC to e.g. H.264/AVC real-time transcoding.

Table 5
DVC potential advantages and current drawbacks for mobile video mail

DVC potential advantages	DVC current drawbacks
Lower encoding complexity	*Lower compression efficiency*
Increased resolution for same power Alternatively to the previous benefit, the resolution of the captured video can be increased while power consumption is maintained. *Improved error resilience* Although this is still a relevant benefit, it is less critical here since in this application scenario parts of the video mail may always be retransmitted.	*Need for a (network) transcoder* *No encoder playback* Editing or playback of captured video at the encoder side is not possible, since it would require a highly complex decoding processing; therefore, mostly non-professional and rather short video mails seem to be possible.

established on the market; in contrast to SMS, MMS may have an arbitrary number of attachments of different types. One possible MMS application is video mail, which can replace SMS in most cases. The benefits of video mail over SMS are obvious: instead of typing, which takes a lot of time, only capturing images and freely speaking is needed with different media replacing difficult textual descriptions of emotions or backgrounds, since 'seeing is believing'.

Table 5 presents the most relevant potential DVC advantages and current DVC drawbacks for the mobile video mail application.

3.6. Disposable video cameras

Disposable cameras appeared in the market first for still pictures and only more recently for video. Disposable or single-use photo cameras are basically a simple box camera sold with a roll of film installed, meant to be used only once. Disposable photo cameras have been around for years and have carved out a healthy niche in the overall photography market. But nobody had come up with a disposable video camcorder until around June 2005 when a $30 one-time-use camcorder went on sale at CVS stores[2] (see Fig. 4, right) [8]. The main business model for

this type of camera revolves around the fact that the device will be used by multiple customers, allowing spreading the cost of the hardware over a number of purchases—at least, if the camcorder is returned to the store for processing. Disposable video cameras are an emerging type of product whose future is still to be seen. It is very likely that more similar products will appear in the market in the near future.

Table 6 presents the most relevant potential DVC advantages and current DVC drawbacks for disposable video cameras.

3.7. Visual sensor networks

With the proliferation of inexpensive cameras (optical sensors) and non-optical (e.g. electrical, thermal, and biological) sensing devices, and the deployment of high-speed, wired/wireless networks, it has become economically and technically feasible to employ a large number of sensing devices for various applications, including embedded devices. Embedded networked sensing may reveal previously unobservable phenomena.

This section is centered on sensor networks using camera sensors. Camera sensor products range from expensive pan-tilt-zoom cameras to high-resolution digital cameras, and from inexpensive webcams and cell phones cameras to even cheaper, tiny cameras such as Cyclops [19]. Due to these advances, the design and

---

[2] CVS Corporation (http://www.cvs.com/CVSApp/cvs/gateway/cvsmain) operates retail drugstores in the United States.

Table 6
DVC potential advantages and current drawbacks for disposable video cameras

DVC potential advantages	DVC current drawbacks
Lower encoding complexity Lower complexity encoding, even if at the cost of some compression efficiency/quality reduction, would be a major plus for this application scenario.	*Lower compression efficiency* *Higher decoding complexity* Although decoding complexity must always lie within reasonable limits, this application may tolerate some higher decoding complexity to 'buy' a reduced encoding complexity since the decoding/transcoding process can be done off-line.
Lower size and weight devices The provision of lower complexity encoders to reach 'low cost', 'low complexity', 'low battery consumption', and 'lightweight devices' is especially important for this application. Although some penalty on the video quality may be acceptable compared to regular video cameras, this penalty should not be too high.	
Flexible allocation of codec complexity	

deployment of camera sensor networks or wireless networks of sensor nodes equipped with cameras are now feasible and useful in a variety of application scenarios. There are many sensor networking applications which can significantly benefit from video information. These applications can include both video-only sensor networks or sensor networking applications in which video-based sensors augment traditional scalar sensor networks. Examples of such applications are security surveillance (civilian or military), environmental monitoring, health care monitoring, and robotics.

In environmental monitoring, a network of wireless camera sensors is used to monitor wild-life habitats or rare species in remote locations. They enable spatially and temporally dense environmental monitoring. Camera sensors can also be used in disaster management scenarios like fire and floods. Since pre-existing infrastructures may be unavailable or destroyed in these settings, a wireless battery powered deployment is necessary. Surveillance so far has been dealing mostly with single stationary cameras, but the recent trend is indeed towards active multi-camera and sensor systems. In particular, the use of multiple video sensors to view a scene is rapidly increasing in many vision-based defense, security, scientific, and commercial applications. These applications may also combine (fuse) images and data coming from other sensors such as optical and infrared sensors, video, Global Positioning System (GPS) and Geographical Information Systems (GIS) data, etc.

In a sensor network, multiple sensors will generate signals which need to be sampled, filtered, transmitted, processed, fused, stored, indexed, and summarized as semantic events to allow efficient and effective queries and mining. Video sensor networks provide a formidable challenge to the underlying infrastructure due to the large computational requirements and the size of the captured data. The amount of video generated can consume the same bandwidth as thousands of scalar sensors. Also, quality and spatial scalability may be relevant issues if the decoder for a specific view uses the image decoded for another view to provide decoded video with increased quality or resolution. In this case, the additional quality or resolution will be provided based on data which is not deterministically known.

Table 7 presents the most relevant potential DVC advantages and current DVC drawbacks for visual sensor networks.

3.8. Networked camcorders

Networked cameras are usually understood as networks of cameras. In this context, networked cameras are typically taken as devices with acquisition, coding, recording, and transmission capabilities, since this is very common in these days. This type of device is also known as 'camcorder' which is a contraction of 'camera' and 'recorder'. The most common application for networks of camcorders is surveillance and monitoring with wired or wireless connections. However, these networks of camcorders are also relevant for shooting and recording in other application contexts like entertainment events such as music concerts, sports, etc. (see Fig. 5). Since there is another section in this paper specifically dedicated to surveillance networks, this section will concentrate on non-surveillance scenarios.

This application scenario is mostly characterized by the usage of multiple devices (cameras/camcorders) for shooting, recording and streaming the same scene, including the capability of later access on demand via wired or wireless channels to the views corresponding to any of the camcorders. This implies for example that the camcorders do not need to be transmitting continuously and simultaneously (they may be accessed one by one depending on the user needs).
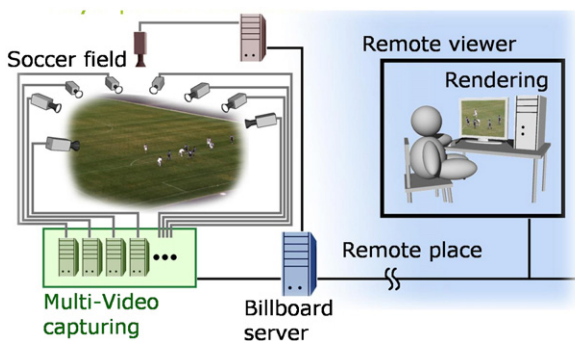
Table 8 presents the most relevant potential DVC advantages and current DVC drawbacks for networked camcorders.

3.9. Distributed video streaming

The huge development of the Internet has given the possibility to realize video streaming systems that allow a user to view a video sequence at his/her own place while receiving it from a remote server or disk. In this setting,

Table 7
DVC potential advantages and current drawbacks for visual sensor networks

DVC potential advantages	DVC current drawbacks
Lower encoder complexity	*Lower compression efficiency* If more complex encoders are not allowed, DVC compression efficiency may still be a drawback.
Lower encoding power consumption In low-power scenarios, the sensor may indeed need to disconnect from time to time. Independent frame encoding as well as data prioritization naturally allowed by DVC coding architectures should facilitate such disconnections and the corresponding decoder re-synchronization.	*Higher decoding complexity* In a dense sensor network, the extra complexity at the decoder may also be seen as a drawback with respect to the scalability of the system this means the number of sensors it can support.
Higher coding efficiency Since most solutions used so far are based on intra coding, e.g. JPEG, DVC solutions may bring here some additional compression efficiency, especially for low encoding complexity. This reduction in the transmission rate with respect to separate encoding and separate decoding is critical for wireless sensor networks since it would allow the use of a higher number of sensors, leading to a better coverage by the sensor network.	
Improved error resilience This advantage is very critical for wireless sensors and may be the most critical for sensor networks in harsh environments.	
Multiview correlation exploitation Inter-view correlation may be a rather important feature in visual sensor networks especially when rather dense visual sensor networks are used.	
Codec/resolution independent scalability	

Fig. 5. Network of cameras shooting a sports event [11]; in this case, the cameras may not have recording capabilities.

the user does not want to download first the video sequence in order to see it at a later time, but he wants instead to see the sequence while it is streaming. With the same idea that led to the development of peer-to-peer networks used for 'distributed' download of files, it is possible to consider the possibility of performing 'distributed streaming' in order to give to the receiver the maximum possible data flow. Here, the video stream is sent to the receiver by various senders in a distributed fashion, in order to reduce the bitrate at the sender sides and increase it at the receiver. In this context, it is possible to consider DVC as a new technology that may be used to perform a more flexible and reliable video streaming system.

Table 9 presents the most relevant potential DVC advantages and current DVC drawbacks for distributed video streaming.

3.10. Multiview video entertainment

Most image and video processing and coding solutions rely on one single camera, referred to as the monoview approach. In the last two decades, extensions to two-camera solutions (also referred to as stereo) have been investigated with limited success in both coding and video analysis applications. Although multiview is also used in solutions with two cameras, here the term will only be used for solutions that use more than two cameras. Multiview video can be used for several applications ranging from free viewpoint television (FTV) to surveillance. In FTV, the user can freely control the viewpoint position of any dynamic real-world scene.

Multiview image and video processing has attracted increasing attention recently and has become one of the potential avenues in future imaging systems, thanks to the reducing cost of cameras. Many tasks can benefit from the availability of multiple views of the same scene, such as interpolation, restoration, segmentation, object recognition, etc. On the other hand, the amount of data captured in multiview video are often tremendous. For instance, in the application of image-based rendering, thousands of images are needed to synthesize novel views from an arbitrary position. This makes data reduction a key issue in multiview image and video processing. Furthermore,

Table 8
DVC potential advantages and current drawbacks for networked camcorders

DVC potential advantages	DVC current drawbacks
Lower encoding complexity	*Lower compression efficiency*
	In this case where the cameras do not communicate, the alternative standards based solution corresponds to the independent encoding of the various views which makes it easier for DVC solutions to also provide advantages in terms of compression efficiency.
Lower size and weight devices	
Lower encoding power consumption	
	Higher decoding complexity
Improved error resilience	
Multiview correlation exploitation	
DVC allows the exploitation of the correlation between different video views, either for the simultaneous transmission of all the views or for the delayed transmission of one view when others have already been transmitted, without requiring the various camcorders to exchange information among them (however, some information like their relative positioning may have to be known). If no easy communication is possible between the various cameras/encoders, DVC may have a definitive advantage regarding predictive codecs.	

Table 9
DVC potential advantages and current drawbacks for distributed video streaming

DVC potential advantages	DVC current drawbacks
Improved resilience and reliability	*Lower compression efficiency*
Using DVC, every sender would provide to the receiver different portions of information without having a precise knowledge of what other senders are doing. This means that in case some of the users disconnect, the system still works as long as sufficient information is globally received from the others. So, with a DVC approach the distributed streaming could be much more flexible to user changes and, for the same reasons, to network faults or rate reallocation.	*Higher decoding complexity*
Flexible allocation of codec complexity	
In case the various senders mentioned above correspond to encoders with different complexity, it may be convenient to have associated decoders with higher complexity to obtain the same decoded quality.	
Codec/resolution independent scalability	

due to the eventual strong correlation between multiple views, multiview data reduction has its own characteristics that differ significantly from traditional image/video compression. As a result, an increasing amount of research on multiview sampling and compression has been proposed in recent years.

Another emerging application field is based on camera arrays (see Fig. 6, left). Large camera arrays can capture multiview images of a scene, which might be used in numerous novel applications such as movie special effects. For camera arrays built for such applications, one of the challenges is the enormous size of the raw data, typically consisting of hundreds of pictures. Hence, compression is needed. To exploit the coherence among neighboring views, the images are usually jointly encoded. In large camera arrays, however, cameras can typically only communicate with a central node, but not between each other. Since joint encoding at the central node requires transmission of all raw images first and excessive memory space to store them temporarily, it is preferable

to compress the images directly at each camera, in a distributed fashion. Existing systems either rely on the built-in compression capabilities at the capturing devices, thus requiring expensive cameras, or need to add customized circuits to perform some form of standard image compression such as JPEG. With hundreds of cameras involved, the cost of either approach may be prohibitive. Multiview video is used in various fields and applications, e.g. high-speed videography, and tele-immersion.

Table 10 presents the most relevant potential DVC advantages and current DVC drawbacks for multiview video entertainment.

3.11. Wireless capsule endoscopy

Many diseases of the human body can only be spotted with images of the ill region. With X-ray, the whole body can be photographed. However, these images are not very
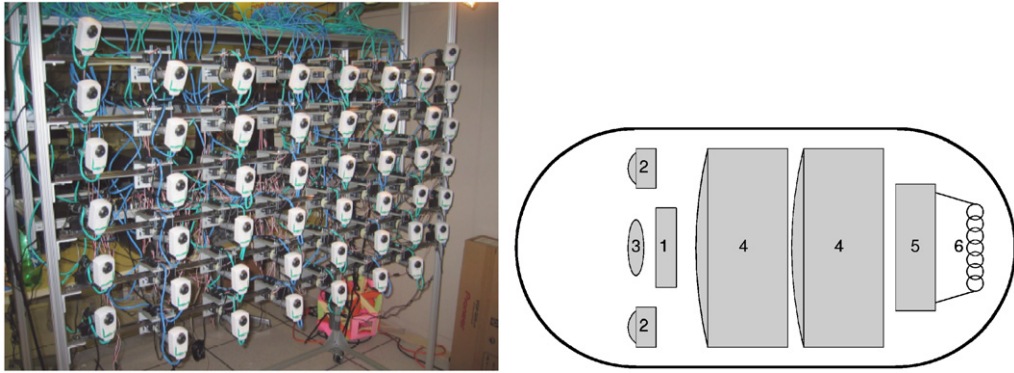
Fig. 6. (Left) Camera array system with 48 cameras [6] and (right) wireless capsule endoscope: (1) CMOS imager; (2) LEDs; (3) lens; (4) batteries; (5) transmitter; (6) antenna [5].

Table 10
DVC potential advantages and current drawbacks for multiview video entertainment

DVC potential advantages	DVC current drawbacks
Lower encoding complexity It has already been shown that DVC encoders provide a significant reduction in complexity when compared to JPEG2000 for large camera arrays compression [25]. *Lower encoding power consumption* *Flexible allocation of codec complexity* *Multiview correlation exploitation* *Higher quality for same complexity* Since there is a trend towards higher quality imaging and, at low bitrates, JPEG2000 tends to blur out image details and incur ringing effects at object boundaries, DVC solutions may be exploited to achieve higher quality instead of reduced complexity.	*Lower compression efficiency* It is important to stress again that the alternative standards based solution corresponds to the independent encoding of the various views which makes it easier for DVC solutions to also provide advantages in terms of compression efficiency. *Higher decoding complexity* *Visual occlusions* For the camera network scenario, it is clear that visual occlusions present a challenging problem for any distributed video coding technique.

accurate, and not all diseases can be detected by this technique. An example is to determine the source of gastrointestinal bleeding. Intestinal bleeding occurs when an abnormality on the inner lining begins to bleed. Determining the source of gastrointestinal bleeding that originates in the small bowel[3] is one of the major diagnostic challenges faced by gastroenterologists. Many small bowel causes of blood loss go undetected because the small bowel is long, hard to reach and therefore difficult to evaluate. X-ray studies may be unable to pinpoint exact locations of abnormalities. Thus, if masses or bleeding lesions are found, their accurate location is difficult to specify to the surgeon for removal. The best way to find most of the causes of small bowel bleeding is to look directly at the small bowel with an endoscope.[4] Since the small bowel is more than 5 m long, which is much longer than any of the instruments currently available, the capsule endoscopy has emerged as an effective way to evaluate the small bowel for bleeding

[22]. The endoscopic capsule has the size of a large pill and contains a battery, a strong light source, a camera and a small transmitter (see Fig. 6, right). Once swallowed, the capsule begins to transmit images of the inside of the esophagus, stomach and small bowel to a receiver. The pictures of the capsule passing through the intestine can be analyzed for abnormalities which are possible reasons for bleeding.

Table 11 presents the most relevant potential DVC advantages and current DVC drawbacks for wireless capsule endoscopy.

4. Characterization and clustering of application scenarios

Following the detailed analysis of application scenarios performed in the previous section, it is possible to cluster the applications according to relevant characteristics such as:

1. number of camera views, single versus multiview systems;
2. real-time performance, real-time versus non-real-time systems;

---

[3] The area of the intestine between the stomach and the colon.
[4] An endoscope is a tube instrument with a light and camera at one end, passed through the mouth.

Table 11
DVC potential advantages and current drawbacks for wireless capsule endoscopy

DVC potential advantages	DVC current drawbacks
Lower encoding power consumption	*Lower compression efficiency*
Lower size and weight devices	*Higher decoding complexity*
Improved error resilience This may be an advantage if transmission and not storage is used.	This drawback may be less critical than usual since availability of powerful decoders in hospitals and clinics might be assumed.

Table 12
Characterization of the studied application scenarios

Application	Multiview	Return channel	Real-time	Error resilience	Power limitations	Encoding complexity	Flexible allocation of codec complexity	Scalability
Wireless video cameras	N	N	Y	Y	Y	Y	Y	N
Wireless low-power surveillance	Y	Y	Y	Y	Y	Y	Y	Y
Mobile document scanner	N	Y	N	Y	Y	Y	Y	N
Video conferencing with mobile devices	N	Y	Y	Y	Y	Y	Y	N
Mobile video mail	N	N	N	N	Y	Y	Y	N
Disposable video cameras	N	N	Y	N	Y	Y	Y	N
Visual sensor networks	Y	Y	Y	Y	Y	Y	Y	Y
Networked camcorders	Y	Y/N	Y	Y	Y	Y	Y	N
Distributed video streaming	N	Y	Y	Y	N	Y	Y	Y
Multiview video entertainment	Y	Y/N	Y	N	N	Y	N	N
Wireless capsule endoscopy	N	N	N	Y	Y	Y	Y	N

3. availability of a return channel, unidirectional versus bidirectional;
4. error resilience, critical versus non-critical;
5. power limitations, low-power constraint versus no low-power constraint;
6. encoding complexity, critical versus non-critical;
7. flexible allocation of codec complexity, critical versus non-critical;
8. scalability, relevant versus non-relevant.

In Table 12, the application scenarios presented in Section 3 are classified in terms of the characteristics listed above. For some cases, it may happen that an application scenario appears both with 'Yes' and 'No' if there are two relevant 'flavors' of that scenario in terms of the characteristic at hand.

Based on the characterization of application scenarios presented in Table 12 and the type of DVC architectures available in the literature (not only most popular but also providing more promising results), four clusters are proposed based on two important dimensions from a DVC point of view: the number of video views, notably monoview versus multiview, and the availability or not of a return channel:

1. applications based on a single view with a return channel available;
2. applications based on multiple views with a return channel available;

3. applications based on a single view without a return channel;
4. applications based on multiple views without a return channel.

These application clusters will be used in the following for the final selection of the DVC most promising application scenarios.

5. Selecting the most promising DVC application scenarios

Following the detailed description and characterization of application scenarios made in previous sections, it is now time to select the most promising application scenarios for DVC by applying the methodology proposed in Section 2, Step 4:

1. For the application scenarios presented in this paper, the matching count between DVC advantages for each scenario and the DVC main benefits previously identified in Section 1 is presented in Table 13.
2. Based on the count shown in Table 13, the market relevance and the target to maximize the coverage of the clusters defined in Section 4 (at least one application per cluster), the application scenarios selected as the most promising ones (in italic in Table 13) are:
   a. distributed video streaming and networked camcorders (bidirectional, monoview cluster);

Table 13

Matching count between DVC application advantages and DVC main benefits

Application scenario	Matching count between DVC application advantages and DVC main benefits
Wireless video cameras	2 (1, 2)
Wireless low-power surveillance	4 (1, 2, 3, 4)
Mobile document scanner	2 (1, 2)
Video conferencing with mobile devices	2 (1, 2)
Mobile video mail	1 (1)
Disposable video cameras	1 (1)
Visual sensor networks	4 (1, 2, 3, 4)
Networked camcorders	3 (1, 2, 4)
Distributed video streaming	3 (1, 2, 3)
Multiview video entertainment	2 (1, 4)
Wireless capsule endoscopy	2 (1, 2)

  b. wireless low-power surveillance (bidirectional, multiview cluster);
  c. wireless video cameras (unidireccional, monoview cluster);
  d. visual sensor networks (unidireccional, multiview cluster).

This selection of application scenarios provides a good balance between monoview and multiview applications as well as between applications with high innovation and applications more established in the market. It is recognized that the simple counting of matches as done above may be simplistic since the various matches do not have necessarily the same weight. However, for the purpose of identifying the most promising applications, this process seems to be enough since too many unknowns are present. While it may happen that DVC will find success in application scenarios not even studied here or no success at all, this is the type of study that can be made at this stage to help moving forward with a clearer understanding of the DVC functional merits.

6. Final remarks

This paper presented in detail application scenarios for which DVC technology may bring major benefits, highlighting the advantages and drawbacks for each application case studied. Based on their detailed description, the application scenarios were after characterized according to some major features, and clustered in four groups of applications. Finally, this paper identified the most promising application scenarios in terms of future exploitation of DVC technology by checking the number of application advantages that match the DVC potential benefits following a proposed methodology.

References

[1] A. Aaron, R. Zhang, B. Girod, Wyner–Ziv coding of motion video, in: Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, November 2002.
[2] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, M. Ouaret, The DISCOVER Codec: architecture, techniques and evaluation, in: Picture Coding Symposium, Lisbon, Portugal, November 2007.
[3] S.C. Draper, A. Khisti, E. Martinian, A. Vetro, J.S. Yedidia, Using distributed source coding to secure fingerprint biometrics, in: International Conference on Acoustics, Speech and Signal Processing, Honolulu, HI, USA, April 2007.
[4] B. Girod, A. Aaron, S. Rane, D. Rebollo Monedero, Distributed video coding, Proc. IEEE 93 (1) (January 2005) 71–83.
[5] F. Hagemüller, M.P. Manns, H.G. Musmann, J.F. Riemann, Medical imaging in gastroenterology and hepatology, in: Series Falk Symposium, vol. 124, December 2002.
[6] ⟨http://amp.ece.cmu.edu/projects/MobileCamArray/⟩.
[7] ⟨http://en.wikipedia.org/wiki/Sousveillance⟩.
[8] ⟨http://www.camcorderinfo.com/content/CVS-One-Time-Use-Video-Camcorder.htm⟩.
[9] ⟨http://www.cylab.cmu.edu/default.aspx?id=2000⟩.
[10] ⟨http://www.homesecuritystore.com/baby_monitor.html⟩.
[11] ⟨http://www.kameda-lab.org/research/project_3dvideo/soccer_3dvideo.jpg⟩.
[12] ⟨http://www.mediaeater.com/cameras/news/taxi.html⟩.
[13] ⟨http://www.videoconferencelighting.com⟩.
[14] Y. Lin, D. Varodayan, B. Girod, Spatial models for localization of image tampering using distributed source codes, in: Picture Coding Symposium, Lisbon, Portugal, November 2007.
[15] M. Maître, Scanner on the Go, Internal IRISA Technical Report, 2005.
[16] A. Majumdar, K. Ramchandran, PRISM: an error-resilient video coding paradigm for wireless networks, in: BROADNETS 2004, San Jose, CA, USA, October 2004, pp. 478–485.
[17] F. Pereira, J. Ascenso, C. Brites, Studying the GOP size impact on the performance of a feedback channel-based Wyner–Ziv video codec, in: IEEE Pacific Rim Symposium on Image Video and Technology, Santiago, Chile, December 2007.
[18] R. Puri, K. Ramchandran, PRISM: a new robust video coding architecture based on distributed compression principles, in: 40th Allerton Conference on Communication, Control and Computing, Monticello, IL, USA, October 2002.
[19] M. Rahimi, R. Baer, J. Warrior, D. Estrin, M. Srivastava, Cyclops: in situ image sensing and interpretation in wireless sensor networks, in: Proceedings of ACM SenSys, Baltimore, MD, USA, November 2004, p. 311.
[20] A.S. Rane, R. Zhang, B. Girod, Wyner–Ziv coding for video: applications to compression and error resilience, in: IEEE Data Compression Conference (DCC), Snowbird, UT, USA, March 2003.
[21] J. Slepian, J. Wolf, Noiseless coding of correlated information sources, IEEE Trans. Inf. Theory 19 (4) (July 1973) 471–480.
[22] M. Tilson, J.R. Saltzman, Small Bowel Bleeding and Capsule Endoscopy, Harvard Medical School, Boston, MA, USA, ⟨http://www.acg.gi.org/patients/gihealth/smallbowel.asp⟩.
[23] A. Wyner, Recent results in the Shannon theory, IEEE Trans. Inf. Theory 20 (1) (January 1974) 2–10.
[24] A. Wyner, J. Ziv, The rate-distortion function for source coding with side information at the decoder, IEEE Trans. Inf. Theory 22 (1) (January 1976) 1–10.
[25] X. Zhu, A. Aaron, B. Girod, Distributed compression for large camera arrays, in: Proceedings of the IEEE Workshop on Statistical Signal Processing, SSP-2003, St Louis, MO, USA, September 2003.