A Comprehensive Overview of Classical and New Perceivable Spatial and Temporal Artifacts in Compressed Video Streams

Dr. Muhammad Uzair,
Faculty of engineering, Islamic University of Madinah, KSA
uzair91@hotmail.com, muzair@iu.edu.sa

Abstract: Modern video encoding tools use lossy coding schemes due to the constraints of power and bandwidth. These coding approaches also create spatial and temporal artifacts which ultimately decrease the received video quality. Understanding these classical and new artifacts, especially produced due to new video encoding tools are very important for the development of efficient video codec's which can avoid these artifacts on the encoder side, rather than compensating them on decoder side. There is still a lack of knowledge about new artifacts due to the new video coding tools, besides the classical artifacts, i.e., blocking, ringing, etc. Many existing papers in the literature describe these artifacts, but none of the paper discusses these artifacts comprehensively. Therefore, this research provides a comprehensive overview of all classical and new artifacts (spatial and temporal) produced by current and new video coding tools, i.e., SVC (scalable video coding) and MVC (multi view coding) by H.264, and NVC (next generation video codec's - H.265/HEVC). The paper also describes the effect of new coding tools on classical artifacts and how these artifacts are related to each other. Overall, this research will assist in the design of more efficient adaptive quantization algorithms and coding mechanisms to improve the video codec performance.

Keywords: Spatial, Temporal, Compressed, Video, Codec's, Artifacts, Blockiness

نظرة عامة شاملة على القطع الأثرية المكانية والزمانية الكلاسيكية والجديدة التي يمكن إدراكها في تدفقات الفيديو المضغوطة

الملخص: تستخدم أدوات تشفير الفيديو الحديثة مخططات تشفير ضياع بسبب قيود القدرة وعرض النطاق. تؤدي أساليب الترميز هذه أيضًا إلى إنشاء نتائج مكانية وزمانية تؤدي في النهاية إلى تقليل جودة الفيديو المستلم. يعد فهم هذه القطع الأثرية الكلاسيكية والجديدة، وخاصة التي يتم إنتاجها بسبب أدوات ترميز الفيديو الجديدة، أمرًا مهمًا للغاية لتطوير برامج ترميز فيديو فعالة يمكنها تجنب هذه القطع الأثرية على جانب التشفير، بدلاً من تعويضها من جانب وحدة فك التشفير. لا يزال هناك نقص في المعرفة حول القطع الأثرية الجديدة بسبب أدوات ترميز الفيديو الجديدة، إلى جانب المصنوعات الكلاسيكية، مثل الحجب، والرنين، وما إلى ذلك. تصف العديد من الأوراق الموجودة في الأدبيات هذه القطع الأثرية، ولكن لم يناقش أي من هذه القطع الأثرية بشكل شامل. لذلك، يقدم هذا البحث نظرة عامة شاملة على جميع القطع الأثرية الكلاسيكية والجديدة (المكانية والزمانية) التي تنتجها أدوات ترميز الفيديو الحالية والجديدة، مثل SVC (ترميز الفيديو القابل للتطوير) وMVC (ترميز متعدد العروض) بواسطة H.264 وNVC (الجيل القادم من برامج ترميز الفيديو -HEVC / H.265.) تصف الورقة أيضًا تأثير أدوات الترميز الجديدة على القطع الأثرية الكلاسيكية وكيف ترتبط هذه القطع الأثرية ببعضها البعض. بشكل عام، سيساعد هذا البحث في تصميم خوارزميات تكميم أكثر كفاءة وآليات تشفير لتحسين أداء ترميز الفيديو.

Introduction: Digital video data gets distorted during acquisition, compression, transmission, decoding and reproduction. The video quality is degraded during the quantization process (lossy compression) to meet the bandwidth, power, and time requirements. Quantization step sizes, especially large, could reduce power consumption, encoding time, and bandwidth requirements, but results in video quality degradation. The perception of the human visual system (HVS) plays an important role during lossy compression as different kinds of spatial and temporal distortions are related to the properties of the HVS. Higher compression ratios are achieved by reducing or eliminating that information which is not noticeable by HVS in order to catch up increasing demand for current industrial video communication system. The quality of the service should be monitored at the receiver side in order to maintain and improve the quality of the received video data, as the demand for a better quality has been higher than ever before. The poor video quality at the consumers end has also resulted in revenue lost in digital communication industry [1].

The compression techniques introduce many visual artifacts (spatial and temporal) in a compressed video. The spatial distortion is visualized in individual frames, while temporal distortion is observed when the video is played. Both spatial and temporal artifacts have many shapes and kinds of distortions [1]. The nature of the artifacts also depends upon which codec is used (MPEG x, H.26x), error position in a frame and whether error concealment is used at decoder or not, etc [2]. In addition to spatial and temporal artifacts, the artifacts produced during acquisition (e.g., camera motion blur), during transmission (e.g., freezing, packet loss), and during video post processing and display (e.g., spatial scaling, chromatic aberration) are also produced. However, these artifacts are not produced during compression [3, 4].

With new coding tools new spatial and temporal artifacts emerged which needs to be addressed, i.e., origin of the artifacts should be known in order to compensate them, etc. Many papers have been written in the past which discusses about spatial and temporal artifacts [3, 4, 5, 6, 7, 8, 9, 10]. However, these papers do not discuss the spatial and temporal artifacts comprehensively. Papers [3, 5, 6] discuss only a few of spatial and temporal artifacts, while paper [7] discuss only the compression artifacts. Similarly, paper [8] discuss only a few of temporal artifact and paper [4] describes only classical spatial and temporal artifacts. Moreover, most of the paper do not discuss new artifacts generated by SVC/MVC or NVC (H.265), except [7]. The existing literature does not discuss all the classical and new artifacts (spatial and temporal) comprehensively in a compressed video. However, this research provides a comprehensive overview and analysis of all spatial and temporal artifacts produced by current and new video coding tools, i.e., MPEG-2, MPEG-4 Part 2, H.264, VC-1, SVC and MVC by H.264, and H.265. The paper also discusses the effect of new coding tools on classical artifacts and explains that how one artifact is related to other artifacts, i.e., masks/creates other artifact, etc. The paper also describes different ways that how these artifacts can be minimized, i.e., need of compensating artifacts at decoder side can be eliminated, etc.

The overall structure of the research is shown in Figure 1.

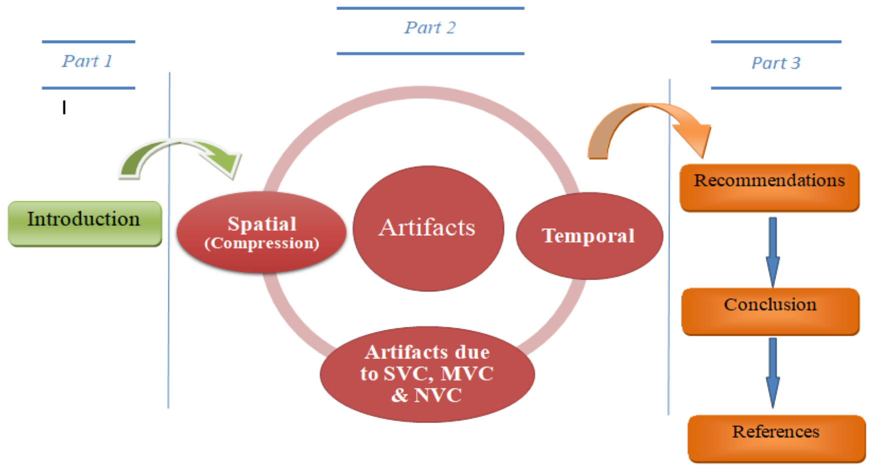

Figure 1: Overall structure of research

Following section presents the spatial artifacts.

1. Spatial (compression based) artifacts

The compression techniques employed in various existing and new video codec's are very similar, i.e., block based DCT (discrete cosine transformation) coding with motion compensation and subsequent quantization. In this kind of coding, quantization errors in the transform domain is the main reason of different kinds of artifacts. The encoding process is also affected indirectly through the quantization scale factor because of motion prediction and decoder buffer size [1]. Different kinds of spatial artifacts are discussed below.

1.1 Blocking artifacts: The blocking artifacts are the most prominent and annoying distortion in block based coding. Blockiness happens because of the two main reasons: 1) independent quantization of individual blocks which leads to discontinuities at the boundaries of adjacent blocks. 2) Due to the motion compensation (MC) prediction, as it reduces inter blocks correlation causing discontinuities near the blocks boundaries. Therefore, using the blocks of different sizes as basic units for transformation, quantization and motion compensation produce discontinuities across block boundaries [8, 9]. The amount and visibility of blocking artifacts increase at low bit rates, i.e., more compressed videos. Blocking artifact is more noticeable in the smoothly textured region, and generally hidden in more spatially active areas (especially when coarse quantization is not used), or bright/dark areas. The lower order DCT coefficient (DC coefficient) is very important to determine whether a blocking distortion will be visible or not [4, 5]. Similarly, natural edges are very abrupt transitions while block edges are weak and regularly spaced. Blocking and slow motion may produce fine-granularity flickering [8].

1.1.1 Effects of macro blocking: For inter predictive coding; macro blocks are partitioned to 16x16, 16x8 and 8x16, down to 4x4 in H.264, H.265. This portioning allows to find better matches for each part by performing a separate motion search for each part [11]. If there is no de-blocking filter, these partitions may increase the blocking and also favor the appearance of motion compensation (MC) mismatch [7].

1.1.2 Effects of spatial scalability: For scalable video coding (SVC) in H.264, up sampling possibly increases blocking artifacts, as the block partitions of the lower layer are up sampled accordingly [7].

1.1.3 Blocking artifacts and H.265/HEVC: In H.264, the macro blocks had a maximum size of 16x16 pixels, and a limited number of sub-blocking patterns. In H.265/HEVC, the macro blocks can be 16x16, 32x32, or up to 64x64 luma sample. The coding efficiency also increases as the macro block size increases. The 64x64 luma size is also called coding tree unit (CTU) which is further divided into prediction units (PU) and transform units (TU) down to a size of 4x4. Due to the CTUs, H.265/HEVC provides new shapes and sizes including all the previous block sizes of H.264. The partitioning does not necessarily change the appearance of artifacts, as smallest size is equal to the H.264. However, blocking artifacts are still obvious, although smaller [2,7,12,13].

1.1.4 Blocking artifacts and different codec's: The blocking artifact appear in all block based video codec's, i.e., MPEG-2 Video, MPEG-4 Part 2, H.261, H.264 and VC-1, etc. MPEG-4 Part 2 and the MPEG-2 coded videos produce blockiness due to its transform size, i.e. an 8x8 DCT. However, the smallest transform size in H.264, H.265, and VC-1 is 4x4 which increases the blocking distortions as the number of block borders increase [7, 11]. Similarly, other block based coding techniques, i.e., vector quantization, block truncation coding, and fractal based compression also suffer from blocking artifact. However, blocking effects are reduced in JPEG 2000, i.e., wavelet transform based compression standards, as the transform is applied to the entire image instead of individual blocks [1].

1.1.5 Filtering blocking artifacts: For the minimization of blocking artifacts, de-blocking is done by the use of simple low pass filtering producing blurring even at those boundaries which do not have blockiness. MPEG-2 Video and MPEG-4 Part 2 perform blockiness compensation at decoder side, as they do not have built in de-blocking filter, while H.264 and VC-1 use an in-loop de-blocking filter to minimize blocking [7, 8]. H.265 only applies the de-blocking filter to 8×8 sample instead of 4*4 as used in H.264, which reduces computational complexity and improves parallel processing operation [2, 12]. In H.264, the reduction of blocking may also produce flickering artifact [8].

1.2 Blurring: Blurring refers to a loss of spatial features and a decrease in edges or texture sharpness due to the removal of high frequency coefficients by quantization or edge-attenuating filters [3]. The coarse quantization further increases blurring. Similarly, de-blocking operators (low pass filters) used by different codec's, i.e., H.264/AVC and HEVC, etc., to reduce blockiness also produces perceptual blurring effect. A video with blurring artifacts due to de-blocking filters appears more pleasant to an observer than without filtering the video [6]. Blurring can also occur due to source related error. In H.264, the reduction of blocking and blurring may produce flickering artifact also [8]. Also, blurring in high spatial activity area may coincide with blocking and mosaic pattern effect as like DCT basis pattern [4].

1.2.1 Blurring and different codec's: The standards, i.e., MPEG-2, MPEG-4 Part 2, H.264 and VC-1 use course quantization which produce blur, and these standards have no integrated filter and use algorithms for de-blurring on the decoder side, if desired. Similarly, producing high frequency components at decoder side (inverse filtering) may yield over sharpening artifacts or introduce noise. The camera motion blur has similar effects like blur, but requires different approach for de-blurring [5, 7].

1.3 Ringing: This artifact is due to the removal of high frequency coefficients because of coarse quantization. It is perceived as ripples and overshoots near high contrast edges/lines and is more prominent in the wavelet coders. It is also known as the Gibbs phenomenon in one dimensional Fourier analysis. It is more visible near sharp and strong edges/lines and those areas where visual masking is weak, i.e., near smooth edges/lines. Both components, i.e., luminance and chrominance, suffer ringing artifacts. It is also not related to the blocking which depends on the existence of uniform/smooth areas, while ringing depends on the amount and strength of edges. A mosquito effect (temporal artifact) is observed by combining ringing with the motion of objects in successive frame [3, 13, 9]. The ringing is more annoying at low bit rate, but it is also observed in low compressed videos [7]. Figure 2 shows ringing artifacts [9].

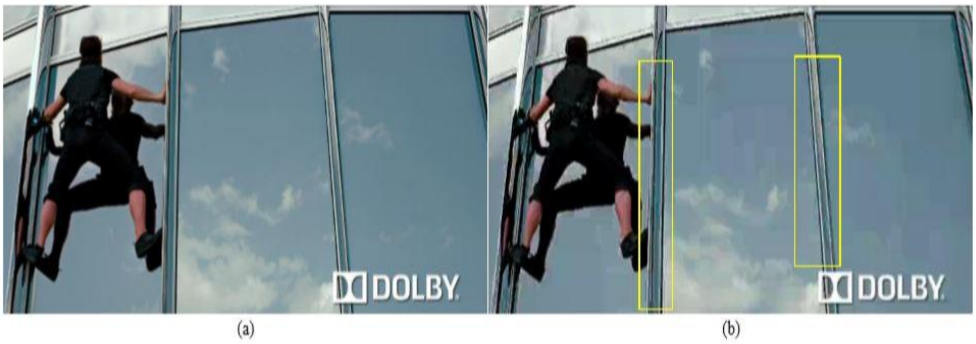

Figure 2. Ringing artifact (a) Reference frame; (b) Compressed frame with ringing artifact [9]

1.3.1 Ringing and different codec's: Ringing is generally reduced with the use of in loop filtering in all video standards, i.e., MPEG-4 Part 2, VC-1, and H.264 and H.265, etc. As 4x4 transform size is used in H.264, VC-1, and H.265, ringing is reduced within one transformed block due to the limited space for over and undershooting [4, 14, 15]. However, the introduction of a 16x16, 32x32 and 64x64 macro block sizes in H.265 may increase ringing artifacts as compared to H.264 codec [2,7, 12].

1.4 Staircase effect: Coarse quantization truncates the higher order basis images to zero and the reconstruction of an image by using lower frequency basis images is either horizontally or vertically oriented, i.e., generally not tuned to the represent diagonal edges/features [5]. Stair case effect typically happens when horizontal and vertical basis functions are not able to accurately represent steep edges due to which horizontal or vertical basis functions becomes more significant. Therefore, a stair case structure is produced along diagonal lines/curve when it get mixed with the false horizontal and vertical edges at block boundaries as shown in Figure 3 (rectangular region) [9]. This artifact is also closely related to ringing and it is more noticeable when the size of a macro block becomes equal to stair case step size. This artifact is also related to blocking and mosaic pattern by demonstrating discontinuities at block boundaries [4, 7].

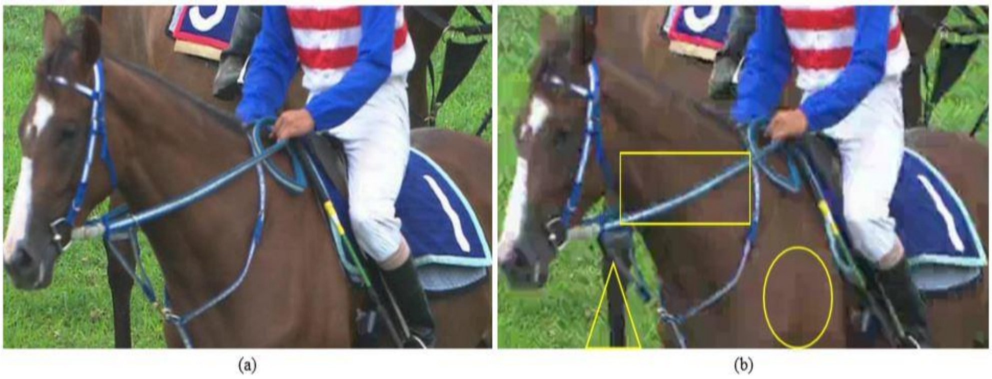

Figure 3. (a) Reference frame (b) Three different types of artifacts:
staircase effect (rectangular area); false edge (triangular area); mosaic pattern (elliptical area) [9]

1.5 False edges: It happens due to the transfer of blockiness distortion from reference to predicted frame due to motion compensation. It produces high frequency noise within the blocks as contrast to blocking which produces same noise along the block boundaries [3]. This effect is more noticeable in smooth areas and generally do not propagate to high spatial activity regions in predictive coded frames [4]. False edge appears near a true edge as shown in Figure 3 (triangular region) [9].

1.6 Mosaic effect (patterns): This effect usually happens when there is mismatch between adjacent blocks, i.e., luminance transitions. Generally, quantization makes all AC coefficients to zero within a block and each block is represented as a constant DC block. Therefore, DC value is different in each block, and by putting them together an abrupt luminance change creates a mosaic pattern. The mosaic pattern is highly visible at smooth regions, i.e., black/white boards, etc. Mosaic pattern is shown in Figure 3 (elliptical region) [9]. Mosaic pattern becomes visually more prominent when two adjacent blocks have different directional orientation. However, mosaic pattern generally reduces if there are many successive frames with high spatial activity area. This effect typically coincides with the blocking effect; however, blackness between two blocks does not always mean the presence of mosaic effect between same two blocks. In case of intra coded blocks, this effect has many similar kind of aspects as basis image effect. Moreover, it may also be introduced by the blocks suffering from the basis image effect [4, 6, 14].

1.7 Color bleeding: Color bleeding happens due to the removal of high frequency coefficients of the chroma components which leads to false color edges. After compression, smearing happens between color channels areas of strongly differing chrominance, as distortions are inconsistent. The resolution of the color channels (Cb and Cr) is half than luminance channel Y and due to the lower resolution interpolation is involved which further enhances inconsistent color spreading resulting in color bleeding [9]. Color bleeding can also happen even without chroma sub sampling. It may also occur due to the incoherent image rendering across the luminance and chrominance channels and can extend over an entire block due to the chroma sub sampling [7, 16]. Figure 4 [10] shows an example of color bleeding (rectangular region) showing chromatic distortion and inconsistent color spreading.

(a) (b)

Figure 4. Color bleeding artifact (a) Reference frame (b) Compressed frame with color bleeding [10]

In chroma sub sampled images, one blurred chroma sample may also extend across multiple luma samples and in such case blurring is also called as color bleeding. Although, chroma sub sampling increases the perceived strength of the color bleeding; color bleeding can also propagate across views, especially using checkerboards arrangements (MVC) [7].

1.8 DCT basis image effect: It happens due to the course quantization when a single DCT coefficient becomes dominant in a block with the reduction of all other coefficients [3]. This artifact has same origin as ringing, but it can also occur other than sharp edges or lines. This effect is also similar to stair case artifacts. Figure 5 shows a basis pattern effect (rectangular regions) [9]. The visibility of the DCT basis pattern also depends on the nature of the texture region (where usually it occurs) [4, 17].

(a) (b)

Figure 5: Basis image effect. (a) Reference frame (b) Compressed frame with basis pattern effect [9]

Due to this effect, when a smooth block is coded it might exhibit strong blocking and blurring artifacts along mosaic pattern effect in adjacent blocks. Moreover, just like mosaic pattern, this effect also decreases over time in high spatial activity area due to the accumulation and refinement of higher order AC coefficients [4, 8].

1.9 False contouring: False contouring can occur as a result of insufficient quantization of DC and low order AC coefficients or their inappropriate distribution. In smoothly texture areas, the pixels change their values gradually and by directly quantizing pixel values in these areas false contouring appears like a step like gradations and may effect to the whole block [4].

1.10 Clipping: It is the truncation of the image values (luminance and chrominance) during arithmetic precision producing abrupt cutting of the peak values at the top and bottom, i.e., creating an aliasing artifacts at peaks as caused by the high frequencies. Peaking, i.e., a sharpness enhancement technique can also produce clipping by adding positive and negative overshoots to the edges. For 8 bit precision, clipping is generally represented as 0 or 255, i.e., percentage of pixels having boundary values [16, 18, 19].

1.11 **Contrast**: It refers as the difference in the luminance value of a pixel of interest and the background. It highly depends on the ability to distinguish an object from its background which is called as dynamic range of a signal. The perception of contrast also varies from human to human as it also depends upon everyone’s mind reference image about objects and sometimes about colors [16, 18, 19].

1.12 Sharpness: It refers to the clarity of details and contours of an image. It can be evaluated using the information provided by edges in the spatial domain or by using high frequencies in the transformed domain. It also highly depends on content, spatial resolution, contrast, and noise [16].

1.13 Noise: It is produced during random processes linked to transmission and generation techniques and is more visible in smooth regions or regions having small variations in the spatial or temporal dimension. Due to this effect, the details and quality of the image degrades.

1.14 Banding effect: Banding effect occurs in large and smooth regions in the reconstructed images when the large quantization step sizes are used. Figure 6 shows the banding effect [20].

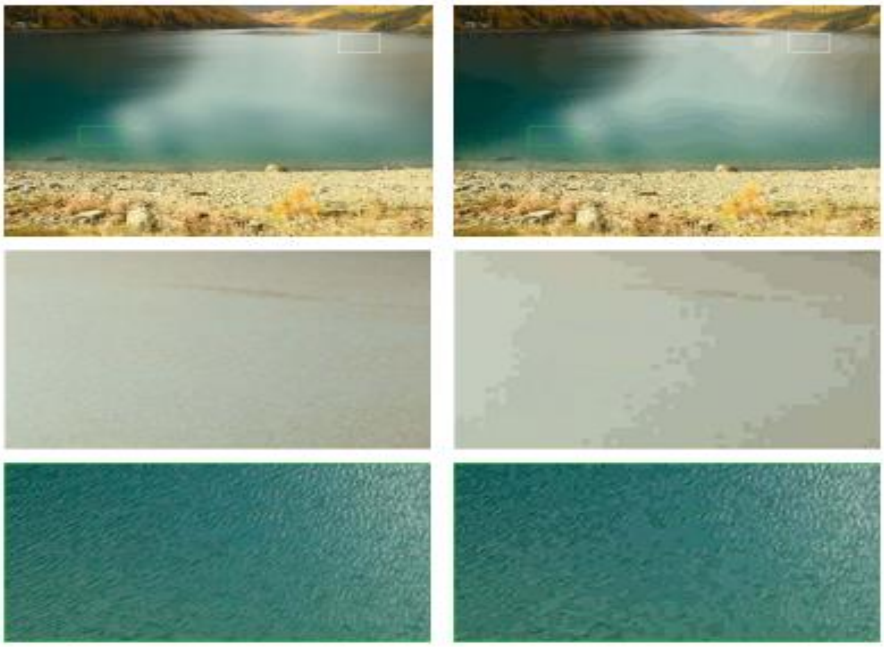

Figure 6: Banding effect (a) Reference frame (b) Restored image with banding effect [20]

The above section has comprehensively discussed the spatial artifacts describing their occurrence reason, relation with other artifacts, and their visual impacts. Figure 7 shows visual a visual pattern of different compression artifacts further highlighting the differences between spatial artifacts.

Figure 7: Different compression artifacts: blocking (marked as 1), blurring (marked as 2), ringing (marked as 3), stair cases (marked as 4), basis patterns (marked as 5), mosaïking (marked as 6) [7]

Table 1 shows a summary of the spatial artifacts [1, 4, 7, 9].

Table 1: Summary of spatial (compressed) artifacts

Artifact	Occurrence reason & Spatial extent within an image	Coexisting artifacts	Relation with other artifacts	Complete removal of a DCT coefficient (truncation effect) or effect of quantization process (association and continuity between blocks)
Blocking	Due to the independent quantization of adjacent blocks; 4*4 blocks up to 64*64 macro block (HEVC)		**1**. Without de-blocking filter, macro block partitioning increases the probability of blocking, and also favors the increase of MC mismatch. **2**. De-blocking filter produces blur. **3**. Blocking & slow motion may produce fine-granularity flickering. **4.** Reduction of the blocking in H.264 may give rise to flickering.	Effect of quantization process
Blurring	Loss of high frequency components; 4*4 block (H.264, H.265)	Ringing at sharp edges; color bleeding (chroma)	**1.** Reduction of the blurring in H.264 may give rise to flickering. **2**. Producing high frequency components at decoder side yield over sharpening artifacts or noise. **3.** Blurring in high spatial activity areas may coincide with blocking and mosaic pattern effect	Truncation effect
Ringing	Insufficient approximation of steep edges; 4*4 block (H.264, H.265)	Blurring	**1.** Mosquito effect is also related to the high frequency distortions as introduced by the ringing effect.	~

Staircase effect	Insufficient approximation of diagonal edge and other features; global spatial extent	Basis image effect for low quantization step sizes	**1**. Relates to blocking and mosaic pattern in terms of highlighting discontinuities between adjacent blocks. **2.** Closely related to ringing and becomes significant when stair case step size equals the size of a macro block through the influence of blocking	Effect of quantization process
False edges	Due to the transfer of blockiness distortion from reference to predicted frame due to MC		**1.** Coincides with the blocking distortion	~
Mosaic Patterns	Apparent mismatch between adjacent blocks due to the truncation of AC coefficients to zero; global spatial extent		**1.** May be produced by the basis image effect **2.** Coincides with blocking effect	Effect of quantization process
Color bleeding	Caused by the coarse quantization of high frequency components of chroma components; spatial extent: 64*64 macro block (HEVC)		**1**.In chroma sub sampled images, blurring is also called as color bleeding	Truncation effect
Basis Image effect	Loss of all but one DCT coefficients; spatial extent: 4*4 block (H.264, H.265)	Stair cases	**1**. May introduce mosaic patterns. **2**. May also introduce blocking and blurring along the boundaries in low spatial activity regions.	Truncation effect
False Contouring	Insufficient quantization of DC and low order AC coefficients or their inappropriate distribution; global spatial extent			Effect of quantization process

*hyphens tell that the distortion does not depend on any other artifact directly or indirectly.

The next section describes the temporal artifacts.

2. Temporal Artifacts

Temporal artifacts are mainly produced when the same region of a frame is coded inconsistently in successive frames of a video sequence. This can happen due to the changes in the type of predictions, quantization levels, motion compensation, or combination of these factors [21, 22]. Temporal artifacts are observed during video playback. It is important to know that many new artifacts have emerged with the advent of new video codec’s as contrast to the spatial artifacts. For an example, texture floating is more significant in H.264/AVC coded video as compared to the early standards. However, it is again minimized comprehensively in the H.265. Similarly, evaluating (objectively) temporal artifact is more difficult as compared to the spatial artifacts. Different kinds of temporal artifacts are described below.

2.1 Flickering: Flickering is a temporal artifact which has significant visual impact. It is basically the change in the luminance or chrominance values along temporal dimension, i.e., coarse quantization from frame to frame. It can also occur due to the variation of prediction techniques from frame to frame, i.e., different intra prediction modes between successive frames. It is mainly observed in static regions where it is more prominent in the background rather than those regions which are in motion. Flickering is generally observed in a low to medium rate coding. Although, blocking and blurring are reduced successfully in H.264, but may increase flickering in intra coded frames [8, 9, 23]. Sometimes, it is also difficult to describe this artifact because of its complexity and appearance variations. Due to the variations, flickering can be further divided into three types based on the appearance and locations, i.e., mosquito noise, fine-granularity and coarse-granularity flickering. Figure 8 shows an example of flickering artifacts [10].

Figure 8. Flickering artifact: Reference frame (left), compressed frame with flickering (right) [10]

2.1.1 Coarse granularity flickering: It refers to the sudden luminance variations in low frequency across Group of Pictures (GOPs. For I-P GOP structure, I frame is intra coded, while all P frames are inter coded using predictions. Between two consecutive GOP's, there is no relation between I frame in the current GOP and the last P frame in the previous GOP. Therefore, a sudden luminance change happens between these two frames (I and P), especially if the scene remains the same. The frequency of this artifact depends on the size of GOP. Using variable GOP lengths and employing a new I frame only when scene change occurs this artifact may be avoided or significantly reduced [1, 8, 9].

2.1.2 Fine granularity flickering: This artifact occurs due to slow motion and blocking distortion in large low to mid energy areas. As considerable blockiness occurs at each frame in low to mid energy regions and the blockiness and DC values also change frame by frame in the corresponding blocks due to the motion or texture details. This causes flashing of these regions at high frequencies, which is eye catching and perceptually annoying [9, 23].

2.1.3 Mosquito noise: Mosquito noise is around sharp object boundaries and is due to both ringing distortions and motion compensation mismatch error. It happens because same region of a frame is coded differently in successive frames. Its appearance is like mosquitoes flying around, i.e., moving along with moving objects, as the plane region has weak visual masking effect and moving objects attract more visual attention. Therefore, this effect is easily noticeable and has huge impact on the video quality. In terms of visibility, ringing distortion has less effect as compared to the motion estimation error. A mosquito like noise can also occur due to the encoder/decoder drift due to the finite precision of the floating point operations in encoding/decoding process. This kind of noise may also be visible at low compressed videos [4, 8, 9, 14, 17].

2.2 Jerkiness: The Jerkiness happens when the motion of the object appears discontinuous. This is due to the reason that speed of the moving object is higher than the available temporal resolution. Jerkiness may become more visible with the motion of strong objects in a frame. Traditionally, jerkiness thought due to the low temporal resolution of video acquisition device, or when some frames are dropped or delayed due to low bandwidth constraint and the video does not remain smooth. In new video coding standards, as frames are divided into layers with coarse to fine temporal resolutions. Due to bandwidth constraints if fine resolution layers need to be dropped, jerkiness might appear. It may also happen due to the transmission delays of the bit stream, i.e., decoder's ability to buffer against fluctuations. Jerkiness can be evaluated using frame rate and temporal activity [4, 5, 9].

2.3 Floating: It refers to the emergence of an illusive motion in certain areas different from their surrounding environment. This illusive movement is wrong as these areas should not move or move together with the background, but they create an illusion due to their movement and look like floating on the top of surrounding background. It happens due to the use of skip mode in video coding, i.e., copying a block in successive frames without updating their details. Figure 9 shows floating artifact [10]. Floating effect can be divided into texture and edge floating depending upon where it appears [9, 10].

Figure 9. Floating artifact: Reference frame (left), Floating artifact (right) [10]

2.3.1 Ghosting (**texture floating):** Ghosting generally happens when a scene having large mid energy textured area (trees or water surface) is captured using slow motion camera, i.e., relative motion between the floating regions and the background creates this effect. It appears like object persistence due to the planned temporal low pass filtering. Rather than actual shifting of image content, video encoder use skip mode for texture regions with zero motion to copy a block from one frame to another (traces of video content remains same in successive frames) in order to save bandwidth and an increase in the mean absolute error. However, this process creates a strong texture floating illusion in an opposite direction at same speed with respect to the camera motion. This effect is typically observed in high energy texture and edge regions. As textures are less visible in very bright or dark regions, therefore, the visibility of this effect is also limited by the luminance levels around the floating areas. This effect is also called texture floating in the literature [4, 5, 9].

2.3.2 Stationary area temporal fluctuations (edge neighborhood floating): This effect most likely also happens due to the use of skip mode and appears at stationary regions next to the boundaries of moving objects. In this case, stationary areas may also appear like a wrapped package surrounding and moving together with the object boundaries rather than remaining stationary. However, this effect may appear without global motion as contrast to the ghosting (texture floating). It is similar kind of fluctuations as associated with the mosquito effect in stationary areas containing major spatial activity, where it is difficult to notice minor differences between a region in one frame and the same region in the next frame of the sequence. This effect is also called as edge neighborhood floating in the literature [4, 7].

2.4 Motion compensation (MC) mismatch: It happens due to the inaccurate motion estimation (ME) during block based matching approach. This results in a mismatched spatial prediction with respect to the current macro block. Motion compensation is generally performed on the luminance value, where the motion vector indicates the spatial displacement of the current macro block from its prediction. Translational block based motion models and motion vectors can cause the inaccurate motion estimation and may produce a reconstructed video with highly visible distortions. In the worst case scenario when the prediction is truncated due to the high error, the reconstruction of the frame will happen with high non-correlation with the current frame. MC mismatch also produces mosquito noise. [24, 25].

2.5 Chrominance mismatch: Chrominance mismatch happens when the same motion vector is used for prediction which has been obtained by using luminance information. In this case, a macro block differs from its own general color and the color of the surrounding area. The chrominance mismatch generally does not appear at object boundaries; therefore, luminance prediction is satisfactory in these areas [3, 25].

2.6 Scene changes: Whenever there is a scene change, a sudden change in spatial features happens in frames before and after the scene. The perceived quality may degrade up to ten percent of the normal resolution after scene cut, although, may restore within half a second. As intra coding is used for the first frame for a new scene, predictive coding is not very efficient in such scenario. The quality of the initial frames is generally poor and builds up gradually as finer spatial characteristics predicatively accumulate. Therefore, there is always a loss of video quality whenever there is a scene change and is generally masked. This degradation in the quality becomes more noticeable when the video is displayed at low rates or as an individual images. One of the solutions is to mix the scenes before and after the cut by preprocessing the source by alpha-mixing. This step provides correlation among frames across the scene change and also avoids the intra frame coding of the first frame after scene change and any kind of related distortions [1, 4].

2.7 Smearing: While recording a video sequence, the light from several moving objects is incorporated at a single point in the recording. Smearing occurs when the recorder simply cannot change the intensity of the beam fast enough to cope with the resolution. Smearing causes loss of spatial resolution and blurring of details. Smearing is visually more noticeable if the viewer is tracking the moving object while watching. The visibility of this effect also depends the speed of the moving object and other spatial masking effects [4, 6].

2.8 Down/up sampling: The up and down sampling of the spatial resolution is not directly related to the block based DCT approach. However, this process is used when the resolution of the source or the display device is not the optimum for coding the algorithm. The down sampling process discards the even fields and makes the vertical resolutions as half. There are many temporal distortions associated with the loss of even field, i.e., changes in vertical sizes, jitter and spatial variations. These artifacts would be generally masked by other spatial and temporal artifacts as they are only visible in small objects or with fine details [4].

2.9 Temporal pumping artifact: H.265/HEVC may use three kinds of prediction, i.e., intra only, low delay (LD) and random access (RA). Temporal flickering artifacts are easily noticeable for intra prediction approach. For LD and RA, the coding efficiency is related to the quantization parameter (QP), which may produce severe quality variations among adjacent pictures in a GOP leading to a perceptually temporal pumping effect (TPA) at medium and low bitrates [26, 27].

2.10 Jitter: This effect relates to the difference in end to end delay between selected packets in a sequence without taking into account any lost packets. This effect can also occur due to the absence of transmission errors creating a similar effect as a lost packet at decoder when the playoff time is missed [3, 16].

2.11 Frame freezing: Frame freezing happens when a single frame is repeated again and again on the screen permanently or often. Frame freezing can also be done on demand to enhance or highlight some thing. Frame freezing deteriorates the received video quality [28].

2.12 Frame skipping: Frame skipping happens when certain frames are not displayed in the video. Sometimes, it is done to improve certain performance, but visual smoothness deteriorates. Frame skipping can even happen due to the hardware failure [29].

2.13 Shimmering Video is encoded in a group of pictures (GOP) in which I frames are intra coded and have high quality, while P and B frames are intermediate frames. Due to the motion estimation error, there can be a sudden drift (difference) between these frames which may increase the P and B frames between successive GOP's. This effect is called the shimmering [16].

2.14 Jagged motion: This effect also happens because of the inaccurate motion estimation. The motion estimation performs well when the movement of all pixels in a macro block is same. However, when the residual error due to motion estimation is large, it is coarsely quantized resulting in jagged motion [3].

2.15 Spread: It is due to the temporal propagation of error between frames due to the motion estimation error. If there is an error in a frame at the start of GOP, then the error spreads due to the prediction from the corrupted frame. It can drift up to whole GOP (until the new I frame in the next GOP), and causes a severe degradation in the reconstructed video.

2.16 Aliasing: When the spatial or temporal contents of a scene are more than the Nyquist rate, aliasing occurs. It can happen between two frames from different scenes, especially when there is a packet loss during scene change. However, the viewer may not notice it when the local movement is slow in the new scene. Temporal aliasing produces the wagon wheel effect, i.e., a situation where an apparent frequency of rotation has been changed by aliasing and a spoked wheel appears to rotate too slowly or even backwards [16, 29].

The above section has comprehensively discussed the temporal artifacts describing their occurrence reason, relation with other artifacts, and their visual impacts. Table 2 shows a summary of the temporal artifacts [1, 4, 7, 9].

Table 2: Summary of temporal artifacts

Artifact	Occurrence reason	Impact	Relation with other artifacts
Flickering	Coarse quantization from frame to frame	Significant impact on perceived video quality	Blocking and blurring are reduced in H.264, but may give rise to flickering
Coarse granularity flickering	Luminance fluctuations in low frequency across Group of Pictures (GOPs).	Frequency and impact depends on the size of GOP	
Fine granularity flickering	Due to slow motion and blocking distortion in low to mid energy regions	Highly noticeable and perceptually annoying	Blocking & slow motion may produce fine-granularity flickering
Mosquito noise	Quantization of high frequency components and MC errors	Strong impact on perceived quality	May also be produced by MC mismatch.

Jerkiness	Speed of the moving object is higher than the available temporal resolution	Become more visible with the motion of strong objects in a frame	
Floating	Encoder uses skip mode to copy a block from one frame to another (encoding of MC predicted residue is skipped)	Appears as an illusive motion in certain areas different from their surrounding environment.	
Ghosting (Texture floating)	Encoder uses skip mode and texture regions with zero motion to copy a block from one frame	More noticeable in high energy texture and edge regions	
Stationary area temporal fluctuations (edge neighborhood floating)	Happens due to the use of skip mode. Occurs at stationary regions next to moving objects.	Stationary areas may also appear like a wrapped package surrounding and moving together with the object boundaries	Similar kind of fluctuations as associated with the mosquito effect in stationary areas
Motion compensation (MC)	Happens due to the inaccurate motion estimation (ME)	High impact on visual quality.	MC mismatch also produces mosquito noise.
Chrominace mismatch	Due to use of same motion vector (obtained using luminance components) for chroma components	A macro block differs from its own general color and the color of the surrounding area	
Scene changes	A sudden change in spatial features in frames before and after the scene, whenever there is a scene change	More noticeable when the video is displayed at low rates or as an individual images.	
Smearing	Occurs when the recorder cannot change the intensity of the beam fast enough to cope with the resolution	Smearing causes loss of spatial resolution and blurring of the details	
Down/up sampling	Happens by discarding even fields during down sampling and making vertical resolutions as half	Causes jittery movement and spatial fluctuations	
Temporal pumping artifact	Happens when quantization varies significantly between adjacent pictures	Severe quality variations among adjacent pictures in a GOP	

The next section presents new artifacts (spatial and temporal) generated due to the use of new coding tools. The section also describes the effect of new coding tools on classical artifacts.

3. New artifacts (spatial and temporal) in new coding tools

Most of the artifacts (spatial and temporal) discussed in section 1 and 2 are present when a video sequence is coded using MPEG-2, MPEG-4 Part 2, H.264, VC-1, H.265. However, the use of new coding tools, i.e., scalable video coding (SVC), multi view coding (MVC), and NVC (next generation video coding) may also introduce new artifacts as described.

3.1 Effect of new transforms and transform sizes in new video standards: MPEG-2 and MPEG-4 Part 2 video standards have an 8x8 DCT transform size. However, the smallest transform size for H.264, H.265 and VC-1 is 4x4. With 4*4 sizes, the blocking distortion may increase as the number of block borders increase. Within one transformed blocks, smaller transform sizes (4*4) also reduce ringing due to the limited space for over and undershooting. H.264 and H.265 can switch between the two integer transform (4x4 and 8x8) while using the High profile. Although, basis functions of the 4x4 DCT and the integer transform are same, but in reality they produce different transform coefficients and transform coefficient distributions for a number of input signals except blocks with DC component only. Integer transform produce less transform coefficients as compared to the DCT avoiding the loss of signal energy during the inverse transform process. There is no current research available to find out that what kind of new artifacts can be produced due to new integer transform used in H.264 and H.265 [2, 7, 11]. H.265/HEVC uses 4*4 transform size, but with DST (discrete sine transform) for only luma blocks, as it better fits with the statistical property. Moreover, 4×4 DST transform is not computationally complex and demanding as compared to the 4×4 DCT transform and also provides approximately 1% bit rate reduction in intra frame coding [1, 2, 7, 12].

3.2 Effect of macro block partitioning: The effect of block partitioning has already been described in blocking and ringing distortion section, i.e., 1.1.1 & 1.3.1, respectively.

3.3 Effect of larger size macro block in H.265: The coding efficiency increases as the macro block size increase. However, the introduction of a 16x16, 32x32 and 64x64 macro block sizes in H.265 enhances the occurrence of ringing artifacts as compared to H.264 codec, which is due to the increased number of coefficients and samples available for over and undershooting [2,7,12].

3.4 Effect of different coding modes: Other than the number and shape of macro block partitions, new artifacts are also produced depending upon which coding mode is used. The flickering or pumping artifact is produced when coding mode of a certain frame area changes in successive frames. This is due to the fact that residuals from prediction differ greatly which produces different coded residual after quantization and produce flickering. The pumping artifacts can be avoided by using similar coding modes for a region in successive frames [2, 7, 11].

3.5 Multi view coding (MVC): In this type of coding, the coding of multiple views of a scene is performed. This is done in order to obtain a three dimensional representation of a scene or part of it. This kind of coding also induces or favors many artifacts. MVC has three different approaches as described below [30, 31].

3.5.1 Depth map quantization: This approach produces these artifacts.

a) Depth ringing (depth bleeding) artifacts: The encoding of a two dimensional image or texture is a type of multi view coding, where depth map indicates the distance of each pixel from the camera. Depth maps are compressed like textures and produce similar artifacts like other codec's [1]. MPEG-4 part 2 explicitly specifies the depth maps coding. The quantization of depth maps yields depth ringing distortions of the depth map [7, 31]. It is also called depth bleeding and it is most noticeable at steep edges of the depth map as shown in Figure 8 [7].

b) Card board or puppet theater effect: Depth estimation error and harsh quantization may also produce card board or puppet theater effect as shown in Figure 10 [7]. As shown in the figure, light colors have more depth as compared to dark colors. The appearance of this artifact is like a two dimensional layers instead of smooth depth transitions, i.e., layer like depth map. The existence of depth map & texture coding may produce a superposition among them and they may mask each other [7, 31].

3.5.2 Frame packing artifacts: Frame packing is a stereoscopic video coding, i.e., second approach of MVC, which is available in H.264 and H.265. In this approach, a single view is used for coding the left and right views using supplemental enhancement information (SEI) message. The interleave coding, i.e., a kind of frame packing approach, of the two views can cause crosstalk of artifacts. The mosquito noise, pumping and MC mismatch may also increase due to the interleaving of views. Side by side and top bottom frame packing can also highlight these artifacts at the borders between the two views. Similarly, column and row alternation as well as checkerboard arrangements can also cause crosstalk. In check board arrangements, color bleeding can also propagate across views. Similarly, using frame alternation provision, MC mismatches increases [7, 31, 32].

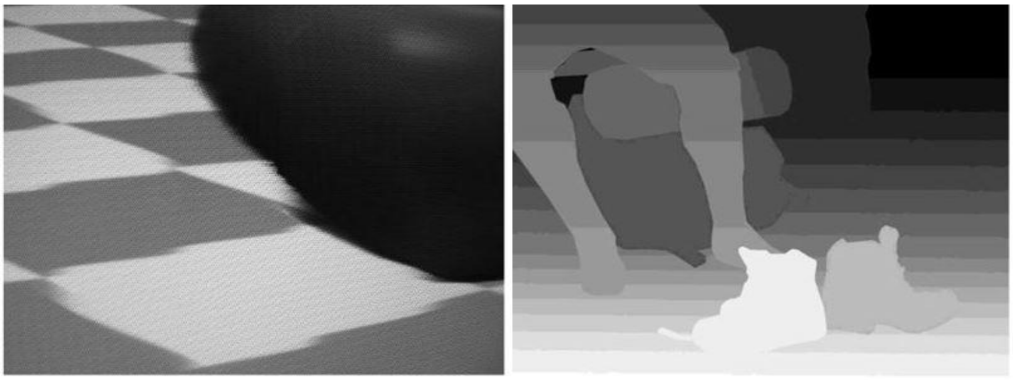

Figure 10: Depth map compression artifacts: a) depth ringing (left), card board effect (right) [7]

3.5.3 Artifacts in H.264 due to MVC: The artifacts can also happen in H.264 due to backwards compatible way of coding (third approach of MVC), i.e., interview prediction, where similarity of many views at any moment is used by MVC. This approach may favor in the increase of MC mismatch artifacts, which needs further investigation [7, 30].

3.6 Scalable video coding: Scalable video coding (SVC) is the encoding of high quality video having one or more sub streams, i.e., layered coding, etc., which is created by dropping packets in order to reduce the bandwidth which is required by the sub stream. Scalable video coding (SVC) refers to decoding required parts of a bit stream to produce smaller frame rate, spatial resolution, or quality [12, 32].

3.6.1 Artifacts due to temporal scalability: Temporal scalability is defined in terms of frame rate. No new artifacts are generated other than mosquito noise, MC mismatch, etc. The quantization parameter is increased in order to avoid pumping artifacts in higher temporal layers [23, 24].

3.6.2 Artifacts due to spatial scalability:

Spatial scalability is defined in terms of spatial resolution. The data which is decoded at lower resolutions can be used to reduce the bit rate and to predict the higher resolutions. Although the basic concept is same, but its implementation is different in MPEG-2 video, MPEG-4 Part 2 and H.264 SVC. In H.264, up sampling possibly increases blocking artifacts, as the block partitions of the lower layer are up sampled accordingly. Up sampling also enhances motion compensation error as same prediction area is used for MC due to the scaling of motion vectors and reusing the reference lists. Moreover, up sampling makes mosquito noise more visible as higher quantization parameters are used in the enhancement layer. Inter-layer intra prediction up sampling may also produce blurring due to bilinear filtering and favors mosquito noise due to the absence of high frequency coefficients [1, 7, 33].

3.6.3 Artifacts due to quality scalability: This scalability is a special case of spatial scalability as the video streams which are generated can be used to predict and decode the video with different qualities. Scalability relies on layers and quality scalability refers to the quality or SNR scalability in terms of reliability or combination of both. Coarse grain quality scalability uses same approach as used by inter layer inter prediction and produce similar artifacts. A drift is also introduced between encoder and decoder if enhancement layers are removed in quality scalability process, which is different from MPEG-2 codec drift, but needs investigation whether it introduces any new artifact or not [1,7, 33].

3.7 Artifacts due to new intra prediction modes in H.265: A new intra prediction mode, i.e., angular intra prediction, has been proposed in H.265 which may increase the probability of pumping artifacts further, which needs further research [2].

3.8 Artifacts due to interpolation filter in H.265: Similarly, in H.265/HEVC, the interpolation filter for subsamples is changed besides the change in transform size, i.e., a 6-tap directional or a 12-tap DCT based interpolation filter, as compared to Wiener and bilinear filter used in H.264. However, signal characteristics are changed in the interpolated subsamples correspondingly, and may expose new artifacts which need further investigation [2, 12].

The above section has comprehensively discussed the new artifacts produced due to the new coding tools and also their effect on classical artifacts. Table 3 shows a summary of these artifacts [1, 4, 7, 9].

Table 3: Summary of new artifacts (spatial & temporal) due to use of new video coding tools

Effect due to new coding tools/Artifact	Impact/Occurrence reason	Relation with other artifacts
Effect of new transforms and transform sizes	1.Integer transform used in H.264 & H.265 produce less transform coefficients as compared to DCT avoiding the loss of signal energy	1.Transform size of 4x4 in H.264/H.265 reduce ringing 3.Smaller transform sizes (4*4) in H.264 & H.265 may increase blocking distortion as the number of block borders increases
Effect of macro block partitioning	1.Enables to perform a separate search for each part of a macro block for fine matching	1.Without de-blocking, partitioning not only increases the probability of blocking but also enhances the appearance of MC (enhanced MC mismatch favors the appearance of mosquito noise)
Effect of larger size macro block in H.265	Coding efficiency increases as the macro block size increase	1. Macro block sizes of 32x32 and 64x64 in H.265 increase the probability of ringing artifacts

Effect of different coding modes	Due to the change of coding modes of one frame area in subsequent frames	Produce flickering or pumping artifact
Depth ringing or depth bleeding (MVC)	Quantization of depth maps yields depth ringing or depth bleeding; most prominent at steep edges	
Card board or puppet theater effect (MVC)	Both depth estimation and harsh quantization may also produce this effect	
Frame packing effect (MVC)	1.Frame packing is stereoscopic video coding (MVC), which can cause cross talk artifacts	1.Color bleeding propagates across views when using checkerboards arrangements, i.e., frame packing. 2.Mosquito noise, pumping and MC mismatch enhances due to interleaving of views. 3.Using frame alternation arrangements, MC mismatches increases
Artifacts in H.264 due to MVC	Artifacts can also happen due to backwards compatible way of coding	1. MC mismatch may increase due to the backwards compatible way of coding, which needs further investigation.
Artifacts due to temporal scalability	Temporal scalability is defined in terms of frame rate.	1. No new artifacts originate other than mosquito noise, MC mismatch, etc. 2. Quantization parameter is increased to avoid pumping artifacts in higher temporal layers. 3. Dropping of fine temporal resolution layers might create jerkiness.
Artifacts due to spatial scalability	Although the basic concept is same, but its implementation is different in MPEG-2, MPEG-4 Part 2 and H.264 SVC	1. Up sampling possibly increases blocking artifacts. 2. Up sampling favors MC mismatches 3. Up sampling makes mosquito noise more visible 4. Up sampling may also produce blurring due to bilinear filtering.
Artifacts due to quality scalability	Quality scalability relies on layers and is defined based on the quality or SNR scalability in terms of reliability or combination	1.Coarse grain quality scalability produce similar classical artifacts
Artifacts due to angular intra prediction mode in H.265	A new angular intra prediction coding mode is used in H.265	May increase pumping artifacts

In the previous sections, we have presented classical and new artifacts (spatial and temporal) comprehensively. Based on the above discussion, the next section proposes recommendations which can help in developing more efficient codec's and dealing with these artifacts more effectively.

4. Recommendations:

4.1 The study has shown that artifacts are generally compensated at the decoder side which may create other visual distortions. However, it would be the best option to develop such codec's which are themselves aware of different kinds of artifacts. In this way, if the sources of these new artifacts are known, then these artifacts can be avoided on encoder side rather than compensating at decoder side. For example, modifying the encoder is best option in order to reduce MC mismatch significantly rather than compensating [1, 7].

4.2 By the development of artifact aware encoder's and also taking into account HVS perception in encoder design, research focus will be on avoiding any artifact rather than compensation. Moreover, this awareness will reduce the post processing issues by the use of information provided by the encoder. In addition, new metrics can be developed which can make artifact detection easier and also apply compensation algorithms more selectively on the encoder and decoder side [1, 7].

4.3 Many codec's perform rate distortion optimization (RDO) to optimize the encoding. HVS effect should be taken into account while performing such calculations [1, 7].

4.4 New video quality algorithms should be developed, especially for most annoying artifacts and for new artifacts generated by new video coding tools.

4.5 New video coding tools (SVC, MVC, and NVC) should be analyzed in detail to find out that how they affect the existing artifacts and can produce new artifacts [2, 12].

4.6 Existing video quality algorithms should be improved in efficiency so that each one of them are quite efficient and reliable to detect classical spatial/temporal and new artifacts

4.7 Pre- and post-processing techniques should also be enhanced and developed to eliminate/minimize the effects of artifacs and enhance the efficiency of image/video quality evaluation.

4.8 Effect of one artifact on other artifacts, i.e., masking/superposition, creation, decrease, etc., needs more in depth analysis, especially the effect of new artifacts on classical artifacts.

4.9 There are many other existing video coding tools whose effects on classical and new artifacts must be analyzed in depth.

4.10 Different artifacts and their effect on HVS need further research as human is final observer of video quality.

4.11 Existing literature generally focus on spatial artifacts. In the same way, temporal artifacts should be analyzed in more detail, especially in new coding tools.

4.12 Video coding standards generally define video decoders only, but encoder configurations are generally different. However, more uniform kind of encoders should be designed to reduce or eliminate the artifacts.

As the blocking distortion is most annoying and it is also present in all classical and new coding tools, therefore, the next section proposes a new novel NR blockiness distortion algorithm for an accurate measurement of blockiness.

6. Conclusion

With the advent of new video coding tools which provide a decreased bit rates compared to previous standards, new spatial and temporal artifacts have emerged along with the classical artifacts. There is still a lack of awareness about classical and new artifacts, especially due to the use of new coding tools. Therefore, this research provides a comprehensive overview of the spatial and temporal artifacts produced by current and new video coding tools, i.e., MPEG-2 Video, MPEG-4 Part 2, H.264, VC-1, SVC (scalable video coding) and MVC (multi view coding) by H.264, and H.265/HEVC. Many existing papers in the literature discuss these artifacts, but none of the paper comprehensively describes all these artifacts as presented in this study. The paper also discusses artifacts produced due to the new coding tools and also their effect on classical artifacts due to these new coding tools.

The first contribution of the paper is a comprehensive survey and analysis of the classical spatial and temporal artifacts as compared to the existing literature. The summary has also been provided for these spatial and temporal artifacts which does not exist in the current literature. The summary further shows that how these artifacts are produced and related to each other, etc. Another contribution of the paper is a comprehensive analysis/survey of the new artifacts (both spatial and temporal) produced due to the use of SVC (scalable video coding) and MVC (multi view coding) by H.264 and H.265/HEVC, respectively. A summary of these artifacts has also been provided which highlights the relation of these artifacts, i.e., how these artifacts are produced and related to other artifacts, etc. Another contribution of the paper is the discussion of the effect of new coding tools on classical artifacts which is not discussed in most of the papers.

Several recommendations have also been proposed as a contribution towards the paper based on the detailed analysis of the artifacts. These recommendations highlights the future work with respect to the knowledge and understanding of the new artifacts produced due the new coding tools, i.e., SVC, MVC, H.265/HEVC, etc., and their effects on the classical artifacts. The artifacts produced due to the new coding tools, especially needs more in depth research.

The research has also several key understandings/observations, i.e., sources of the existing and new artifacts should be known, etc. The artifacts aware encoders can be developed with this knowledge which is more efficient rather than compensating the artifacts at the decoder side. This research can be used in any image processing applications where received quality of the image/video is not up to the standard and needs further investigation, especially to see the effect of one artifact to the others.

7. References

[1] Richardson, I.E.G. (2003) H.264 and MPEG-4 Video Compression. Chichester, England: Wiley & Sons Ltd.

[2] Sullivan, G., Ohm, J., Woo-Jin, H., and Wiegand, T., “Overview of the High Efficiency Video Coding (HEVC) standard,” IEEE Trans. Circuits and Systems for Video Technology 22(12), 1649–1668 (2012).

[3] B. M. Zavodsky, “Time-variant video quality evaluation for mobile networks,” Phd Thesis, Fakulta elektrotechniky a informatiky Katedra rdioelektroniky, Slovensk technick univerzita v Bratislave, 2006.

[4] Yuen, M. and Wu, H., “A survey of hybrid MC/DPCM/DCT video coding distortions,” Signal Processing 70 (3), 247–278 (1998).

[5] WeisiLin, C. Jay Kuo, “Perceptual visual quality metrics: A survey”, Journal of Visual Communication and Image Representation, volume 22, Issue 4, May 2011, pages 297-312.

[6] Muhammad Shahid, Andreas Rossholm, Benny Lövström, and Hans Zepernick, “No-reference image and video quality assessment : a classification and review of recent approaches”, EURASIP Journal on Image and Video Processing, 2014, pages 397-437.

[7] Andreas Unterweger, “Compression artifacts in modern video coding and state of the art means of compensation”, University of Salzburg, Austria, 2012.

[8] Chodisetti Rakesh Anil, “H.264 video coding artifacts; Measurement and reduction of flickering ,” Master thesis, Department of Electrical Engineering, Blekinge Institute of Technology, 2013.

[9] K. Zeng, T. Zhao, A. Rehman, and Z. Wang, “Characterizing perceptual artifacts in compressed video streams,” IS&T/SPIE Human Vision and Electronic Imaging, vol. 9014, no. 10, pp. 2458–2462, Feb. 2014.

[10] Liqun Lin, Shiqi Yu, Tiesong Zhao, Member,” PEA265: Perceptual Assessment of Video Compression Artifacts”, Computer Vision and Pattern Recognition; Image and Video Processing, 1601–1612 (2019).

[11] Zeng, K., Rehman, A., Wang, J., and Wang, Z., “From H.264 to HEVC: Coding gain predicted by objective video quality assessment models,” in International Workshop on Video Processing and Quality Metrics for Consumer Electronics, Feb 2013.

[12] Hanwha Techwin, “White Paper: H.265 High Efficiency Coding Video Compression for Security Applications”, Ridgefield Park, NJ, USA, 2016

[13] K. Zhu, C. Li, V. Asari, and D. Saupe, “No-reference video quality assessment based on artifact measurement and statistical analysis,” IEEE Trans. Circuits Syst. Video Technol., vol. 25, no. 4, pp. 533–546, April, 2015.

[14] S. Yoo, K. Choi, and J. Ra, "Blind post-processing for ringing and mosquito artifact reduction in coded videos," IEEE Trans. Circuits Syst. Video Technol., vol. 24, no. 5, pp. 721–732, May 2014.
[15] H Liu, N Klomp, I Heynderickx, "A no-reference metric for perceived ringing artifacts in images", IEEE Trans. Circuits Syst. video technol. volume 4, issue 20, pages 529–539, October, 2010.
[16] N. Suresh, "Mean time between visible artifacts in visual communications," Phd Thesis, Department of Electrical and Computer Engineering, Georgia tech. University, 2007.
[17] C. Dong, Y. Deng, C. Chang, and X. Tang, "Compression artifacts reduction by a deep convolution network," in Proc. IEEE International Conference on Computer Vision (ICCV), Feb. 2016, pp. 576–584.
[18] J. Xia, Y. Shi, K. Teunissen, and I. Heynderickx, "Perceivable artifacts in compressed video and their relation to video quality," Signal Process Image., vol. 24, no. 7, pp. 548–556, Aug. 2009.
[19] M. Rohil, N. Gupta, P. Yadav, "An improved model for no-reference image quality assessment and a no-reference video quality assessment model based on frame analysis", Signal, Image and Video Processing, vol 6, no.7, pp. 610-619, July 2019.
[20] J. Liu , D. Liu, W. Yang , S. Xia , X. Zhang , and Y. Dai," A Comprehensive Benchmark for Single Image Compression Artifact Reduction", IEEE Transactions on image processing, vol. 29, pp. 225-241, 2020
[21] K. Yang ; S. Wan ; Y. Gong ; Y. Yang ; Y. Feng, "Temporal Distortions in H.265/HEVC", Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), Jeju, South Korea, 13-16 Dec. 2016.
[22] Y. C. Gong, S. Wan, K. Yang, F. Z. Yang, and L. Cui, "An efficient algorithm to eliminate temporal pumping artifact in video coding with hierarchical prediction structure," J. visual communication, vol. 25, no. 7, pp. 1528–1542, Oct. 2014.
[23] Y. Yang, J. Xiang. and H. Wau., "Robust filtering technique for reduction of temporal fluctuation in H.264 video sequences," IEEE Trans. Circuits and Systems for Video Technology 20(3), 458–462, Nov, 2010.
[24] J. Lee, J. Park, H. Wang, "A New Distortion Measure for Motion Estimation in Motion-Compensated Hybrid Video Coding. Signal Processing", Image Communication, vol. 2, no. 6, pp. 75–84, 2011.
[25] P. V. Pahalawatta and A. M. Tourapis, "Motion estimated temporal consistency metrics for objective video quality assessment," in Quality of Multimedia Experience, International Workshop on, pp. 174–179, 2009.
[26] S. Wan, Y. Gong, and F. Yang, "Perception of Temporal Pumping Artifact in Video Coding with the Hierarchical Prediction Structure," in Proc. IEEE Int. Conf. Multimedia and Expo, pp. 503-508, Jul. 2012.
[27] Y. Gong, S. Wan, K. Yang, B. Li, and H. Wu, "Perception based Quantitative Definition of Temporal Pumping Artifact", in Proc. of the 19th Int. Conf. on Digital Signal Processing, pp.870-875, August 2014.

[28] Y. Qi and M. Dai , “The Effect of Frame freezing and Frame skipping on video quality”, Proceedings of the 2006 International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Jakarta, Indonesia, 13-16 April, 2006.

[29] Z. Akhtar and T. Falk, “Audio-Visual Multimedia Quality Assessment: A Comprehensive Survey”, IEEE access, vol. 11, no. 7, pp. 412-440, July, 2017

[30] G. Vetro, A. Wiegand, T. Sullivan, “Overview of the Stereo and Multiview Video Coding Extensions of the H.264/AVC Standard”. Proceedings of IEEE, Special Issue on 3D Media and Displays, vol. 99, no. 4, pp. 626–642, 2011.

[31] P. Merkle, Y. Smolic, P. Wiegand, “The effects of multiview depth video compression on multiview rendering”, Signal Processing: Image Communication, vol. 24, no. 1, pp. 73–88, 2008.

[32] H. Boujut, B. Pineau, J. Hadar, O. Ahmed, “Weighted-MSE based on saliency map for assessing video quality of H.264 video streams”, Proceedings of the SPIE “Image Quality and System Performance VIII”, vol. 7, no. 6, pp. 686–696. 2011.

[33] R. Farrugia, A. Debono,”A Hybrid Error Control and Artifact Detection Mechanism for Robust Decoding of H.264/AVC Video Sequences”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 5, pp. 756–762, 2010.