Maximizing Image Quality over Visual Sensor Networks via DCT Bit Allocation

Sungjin Lee, Sanghoon Lee
Electrical and Electronic Engineering,
Yonsei University, Seoul, Korea
elflee7@yonsei.ac.kr, slee@yonsei.ac.kr

Alan C. Bovik
Department of Electrical and Computer Engineering
The University of Texas at Austin, Austin, TX, USA
bovik@ece.utexas.edu

***Abstract*—Recently, VSNs (Visual Sensor Networks) are becoming of increased interest in a variety of applications. Due to power constraints at each visual sensor node, the processing of live video is constrained. In this paper, we propose a DCT (Discrete Cosine Transform)-based bit allocation scheme to maximize image quality for a given bit rate constraint. Two major featured algorithms are proposed : "SLO" (Searching from the Lowest Order) and "SPP" (Searching the Present optimal bit allocation from the Previous optimal bit allocation). The proposed algorithms reduce the search range based on a simple IQA (Image Quality Assessment) model and DCT statistics. Thus, they demonstrate low-complexity and fast computation. In the simulations, we show the superiority of the proposed algorithms over conventional approaches.**

***Keywords*-DCT coefficient, Image Quality Assessment, SSIM, Visual Sensor Network, Energy-Limited Transmission**

I. INTRODUCTION

Recently, VSNs (Visual Sensor Networks) driven by the availability of inexpensive CMOS cameras, has drawn increasing attention for applications such as environmental monitoring and ad-hoc surveillance [1][2]. However, in wireless applications, due to the limited energy available to deal with the large volume of multimedia data, image processing is done at each sensor node. In order to realize efficient image processing with a limited bit budget, it is necessary to develop a low-complexity algorithms that conserve internal processing power while maintaining high visual quality.

To alleviate this problem, a few papers propose compression techniques using spatial and temporal samplings [5][6], reducing spatial correlation between cameras. However, such compression schemes are available only for very densely deployed networks. Moreover, the compression techniques used are commonly designed by measuring the reconstructed image quality using the PSNR (Peak Signal to Noise Ratio), which correlates poorly with subjective image quality.

Here, we consider spatial image compression assuming reduced inter-camera correlation over the VSN, using DCT-based unequal-rate quantization to maximize the image quality (as measured by SSIM) of each captured scene. In particular, we design a novel algorithm that seeks an optimal quantization-rate for a given bit-rate. Since the proposed algorithm reduces the search range based on a simple quality model and simple DCT statistics, it has low-complexity.

II. QUALITY ASSESSMENT MODEL

We define the following parameters of the captured images and the quantization process.

-$s[t]$: Image sensed at time slot t,

-$q[t]$: Set of quantization parameters at time slot t,

-$\omega[t]$: Quantized image of $s[t]$ by $q[t]$

-b: Index of MB (Macro Block) in image $s[t]$,

-B: Number of MBs in image $s[t]$,

-p: Index of pixel in MB b,

-P: Number of pixels in MB b, which is 8×8 for DCT,

-$v[b][p]$: Quantization rate applied to pixel p in MB b,

-$V[b]$: Block set of $v[b][p]$, $V[b] := \{v[b][p]\}$

-$x[b][p]$: Value of pixel p in MB b,

-$X[b]$: Set of $x[b][p]$, i.e. $X[b] := \{x[b][p]\}$

-$\Theta[b]$: DCT coefficient set of $X[b]$,

-$\theta[b][p]$: Value in pixel p of $\Theta[b]$. These values are ascendantly ordered according to the frequency, i.e. $\theta[b][0]$ is the DC value of $\Theta[b]$.

Further, the following parameters of the quantized image are defined by

-$y[b][p]$: Pixel value of quantized image $\omega[t]$,

-$Y[b]$: Set of pixel values of quantized image $\omega[t]$,

-$\Lambda[b]$: DCT coefficient set of $Y[b]$,

-$\lambda[b][p]$: Value in pixel p of $\Lambda[b]$, $\Lambda[b] := \{\lambda[b][p]\}$.

For brevity, we denote X, x, Y, y instead of $X[b]$, $x[b][p]$, $Y[b]$, $y[b][p]$.

In image coding, it is difficult to perform optimal bit allocation using an IQA index such as SSIM (Structural SIMilarity) [4]. Thus, we introduce a heuristic based search algorithm for optimal bit allocation using the DCT coefficients of space domain vectors following the results in [3][4]. In particular, the relation between the SSIM index and the quantization rate is presented.

The SSIM index can be written

$$\begin{aligned}
&\mathrm{SSIM}(X[b], Y[b]) \\
&= [l(X,Y)]^{\alpha}[c(X,Y)]^{\beta}[s(X,Y)]^{\gamma}, && (1)\\
&= \left[\frac{2\mu_x\mu_y + C_1}{\mu_x^2+\mu_y^2+C_1}\right]\left[\frac{2\sigma_{xy}+C_2}{\sigma_x^2+\sigma_y^2+C_2}\right], && (2)\\
&= \left[\frac{2\frac{\theta[b][0]\lambda[b][0]}{P}+C_1}{\frac{\theta[b][0]^2+\lambda[b][0]^2}{P}+C_1}\right]\left[\frac{2\frac{\sum_{p=1}^{P-1}\theta[b][p]\lambda[b][p]}{P-1}+C_2}{\frac{\sum_{p=1}^{P-1}\theta[b][p]^2+\lambda[b][p]^2}{P-1}+C_2}\right]. && (3)
\end{aligned}$$

In Eq. (1), the terms $l(X,Y) = (2\mu_x\mu_y + C_1)/(\mu_x^2 + \mu_y^2 + C_1)$, $c(X,Y) = (2\sigma_x\sigma_y + C_2)/(\sigma_x^2 + \sigma_y^2 + C_2)$, and $s(X,Y) = (\sigma_{xy} + C_3)/(\sigma_x\sigma_y + C_1)$ are used to obtain the luminance, contrast and structural correlation of the signals, respectively. Under the assumption $\alpha = \beta = \gamma = 1$, and $C_3 = C_2/2$, Eq. (2) can be derived from Eq. (1). In addition, by substituting the space domain mean, variance, and cross correlation terms in Eq. (2), Eq. (3) is derived.

Hence, the average SSIM index $\overline{\mathrm{SSIM}}(s[t], q[t])$ over time slot t is

$$\overline{\mathrm{SSIM}}(s[t], q[t]) = \frac{1}{B}\sum_{b=0}^{B-1}\mathrm{SSIM}(X[b], Y[b]), \qquad (4)$$

III. Optimization Formulation

In this section, we present an algorithm to allocate a given bit-rate into each DCT group in a picture to maximize the SSIM score. We divide DCT coefficients into a few groups, rather than a fixed block size of 8×8, for faster implementation.

Unfortunately, since SSIM is not concave, it is not possible to employ a regular convex optimization approach for bit allocation. We resolve the problem by formulating it as an integer programming optimization problem. Let $R_{k,t}$ be the rate $R_{k,t}$ for DCT group k :

$$\mathbf{R}_t := \{R_{k,t}\}, \quad k \ \text{ integer} \in [1, N_G],$$

where $R_{k,t} = v[b][(k-1)\cdot M_G + n],\ n \in [0, M_G - 1],\ b \in [1, B]$. Here, k is the index of the DCT group, N_G is the number of DCT groups in an MB, $\mathbf{R}_t$ is the set of rates $R_{k,t}$ after quantization and M_G is the number of elements in a DCT group. When M_G is assumed to be equal for each group, N_G and M_G should be a power of 2. For example, when $P = 8 \times 8$, $M_G = 16$ and $N_G = 4$, the parameters $\mathbf{R}_t$, $R_{k,t}$ become

$$\begin{aligned}
&\mathbf{R}_t := \{R_{k,t}\}, \quad k = (1,2,3,4),\\
&R_{k,t} = v[b][(k-1)\cdot 16 + n], \quad n \in [0, 15], \quad b \in [1, B].
\end{aligned}$$

When assuming a DCT size of 8×8, then N_G can be theoretically $2^0,\ 2^1,\ 2^2,\ \ldots,\ 2^6$. However, we will consider the case of 2^2 for ease of understanding, which also can be applied into the other cases.

Then, employing the SSIM function in terms of $\mathbf{R}_t$, we formulate the problem as follows:

$$\begin{aligned}
\mathbf{R}_t^* \quad &= \quad \arg\max \overline{\mathrm{SSIM}}(s[t], \mathbf{R}_t) && (5)\\
\text{subject to} \quad & \quad \mathbf{1}^T\mathbf{R}_t \le C[t],\\
& \quad 1 \le R_{k,t} \le 8,\\
& \quad R_{k,t} \text{ integer}, \ (k = 1, \ldots, N_G).
\end{aligned}$$

where the control parameter $R_{k,t}$ should be integer with a range from 1 to 8, and the sum of $R_{k,t}$ over domain k should be less than a given bit-rate $C[t]$. Before describing this algorithm, we discuss some features from an example of rate allocation.

In this example, we assume that N_G is 4, the size of a sample image is 512×512, the size of a DCT block is 8×8. The sample images considered are “Barbara”, “Lena” and “Mandrill”. In addition, we introduce a new parameter “order” to analyze the level of preference for low frequencies. For the practical cases of bit allocation $\mathbf{R}_t$, with $C[t] = 5$, what is “they” should be $(2,1,1,1)$, $(1,2,1,1)$, $(1,1,2,1)$ and $(1,1,1,2)$. Here, the heaviest allocation on low frequencies is $(2,1,1,1)$ which is defined as the first order among the total four cases, $(1,2,1,1)$ is defined as the second order, and so on. In other words, the lowest orders heavily favor low frequencies. If the optimal allocation is performed at the lower orders, a low order allocation is more beneficial than other allocations, because it better reflects the distribution of frequencies.

Table I shows a similar tendency for $C[t] \le 8$, which show that the lowest order maximizes the SSIM score. On the other hand, as $C[t]$ increases ($C[t] \ge 12$), the order increases as well, i.e. the numbers of bits in the other blocks, ($R_{2,t}$, $R_{3,t}$, $R_{4,t}$) increase, since bit allocation to the first block is saturated. Therefore, it is reasonable to search the available cases starting from the lowest order over a reasonable range. The range can depend on the frequency distribution of each image, but it is certain that the scheme is much faster than simple full search.

The other important feature to consider is correlation between the bit allocations for $C[t]$ and $C[t] + 1$. For example, $\mathbf{R}_t$ for $C[t] = 7$ and $C[t] = 8$ in the image “Barbara” are $(4,1,1,1)$ and $(4,2,1,1)$. Here, the only difference is the number of bits in the second block. This concept is observed at the other values of $C[t]$, i.e. the present optimal bit allocation can be easily deduced from the previous optimal bit allocation using at most four searches. In this way, we can find the optimal bit allocations from $C[t] = 5$ to $C[t] = 15$ with a search computation of only $44\ (= N_G \times 11)$ using the initial allocation $(1,1,1,1)$.

The application domains of these two ideas may differ. The first feature “*SLO*” (Searching from the Lowest Order) can be more advantageous for the one-time search problem and for images greater low frequency energy. On the other hand, the second idea “*SPP*” (Searching the Present optimal

Table I
OPTIMAL RATE ALLOCATION USING SSIM SCORE FOR SAMPLE IMAGES

	Barbara			Lena			mandrill		
$C[t]$	$\mathbf{R}_t$	order	SSIM	$\mathbf{R}_t$	order	SSIM	$\mathbf{R}_t$	order	SSIM
5	(2,1,1,1)	1 / 4	0.466	(2,1,1,1)	1 / 4	0.564	(2,1,1,1)	1 / 4	0.556
6	(3,1,1,1)	1 / 10	0.550	(3,1,1,1)	1 / 10	0.671	(3,1,1,1)	1 / 10	0.632
7	(4,1,1,1)	1 / 20	0.617	(4,1,1,1)	1 / 20	0.763	(3,2,1,1)	2 / 20	0.690
8	(4,2,1,1)	2 / 35	0.678	(5,1,1,1)	1 / 35	0.821	(4,2,1,1)	2 / 35	0.729
9	(5,2,1,1)	2 / 56	0.732	(6,1,1,1)	1 / 56	0.854	(4,3,1,1)	5 / 56	0.767
10	(5,3,1,1)	5 / 84	0.786	(6,2,1,1)	2 / 84	0.890	(4,3,2,1)	12 / 84	0.805
11	(6,3,1,1)	5 / 120	0.817	(6,3,1,1)	5 / 120	0.913	(4,3,3,1)	24 / 120	0.830
12	(6,4,1,1)	10 / 161	0.850	(7,3,1,1)	4 / 161	0.930	(5,3,3,1)	23 / 161	0.853
13	(6,5,1,1)	17 / 204	0.871	(7,4,1,1)	7 / 204	0.943	(5,3,3,2)	39 / 204	0.876

bit allocation from the Previous optimal bit allocation) can be profitable for the multiple search problem and for images containing more high frequency energy. By employing these two features, we design the algorithm as follows,

`if` $(\mathrm{NPS} < 0.2)\ \&\ (\mathrm{NCS} \leq 3)$
    `do` *SLO algorithm*
`else`
    `do` *SPP algorithm*
`end`

where NPS is a function that captures the percentage of image energy at low frequencies. A simple measure is defined as the number of suprathreshold DCT coefficients with high magnitude.

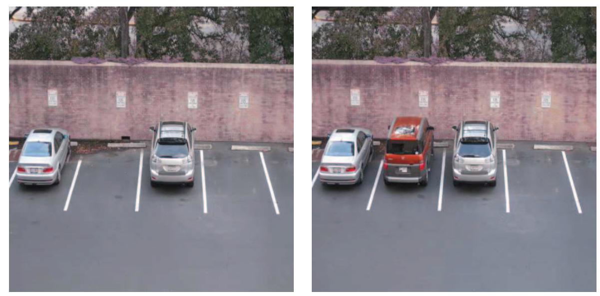
$$\mathrm{NPS} = \frac{\mathcal{N}(\ \mathrm{abs}(\theta[b][p]) > 10\ ,\ \forall b, \forall p)}{|s[t]|}, \tag{6}$$

where $\mathcal{N}$(condition) is cardinality operator. The other quality, NCS represents the number bit rates to search. For example, when $C[t] = 5$, $C[t] = 6$ or $C[t] = 7$, NCS becomes 3.

In addition, the *SLO algorithm* is

`For` index $= [1 : \mathrm{NCS}]$
    $C[t] = \mathrm{SumBit}(\mathrm{index})$,
    `make` $\mathrm{OB}(\mathbf{R}_t)$ `using constraints in` (5),
    `find` $\mathbf{R}_t^*$, i^* `that maximizes`
        $\overline{\mathrm{SSIM}}(s[t], \mathrm{OB}(\mathbf{R}_t, i)),\quad i \in [1,\ st \cdot |\mathrm{OB}(\mathbf{R}_t)|\ ]$
`end`

where $\mathrm{OB}(\mathbf{R}_t)$ is the ordered bit combinations assuming preferential bit allocation toward lower orders. In particular, $\mathrm{OB}(\mathbf{R}_t, i)$ is the i^{th} bit combination in $\mathrm{OB}(\mathbf{R}_t)$. In addition, st is a statistical factor that determines the search range, which increases the efficiency of the optimal solution by utilizing the statistical pattern about the optimal rate allocation of the captured image. In general, when $st = \mathrm{NPS}$, it is well matched. SumBit is the set of sum bits to search in ascending order. $\mathrm{SumBit}(i)$ is the i^{th} value among SumBit.



(a) image at τ_0 (b) image at τ_0+$\Delta\tau$

Figure 1. Captured images at τ_0 and τ_0+$\Delta\tau$

The *SPP algorithm* is

`Init` $R_{k,t}^* := \{1\}$ `at` $C[t] = N_G,\ \forall k \in [1, N_G]$,
`For` $C[t] = [N_G + 1 : \mathrm{UC}]$
    `Find` $\mathbf{R}_t^*,\ k^*$ `at` $C[t]$ `that maximizes`
        $\overline{\mathrm{SSIM}}(s[t],\ \mathcal{A}(\mathbf{R}_t, C[t], k)),\ \forall k \in [1, N_G]$
`end`

where UC is the maximum given bit-rate, and $\mathcal{A}(\mathbf{R}_t, C[t], k)$ represents $\mathbf{R}_t$ at $C[t]$ by adding 1 bit to the k^{th} block in $\mathbf{R}_t^*$ at $C[t] - 1$.

This algorithm significantly reduces the computation.

IV. Simulation Results

To simulate the proposed algorithm, we use images of a parking surveillance as shown in Fig.1, captured at times τ_0 and $\tau_0 + \Delta\tau$. In the simulation, we assume that $N_G = 4$, the size of images is 512×512 and the size of DCT block is 8×8.

Table II shows the optimal rate allocation of the proposed algorithm at each time. From these results, it is apparent that the results between τ_0 and $\tau_0 + \Delta\tau$ are very similar. It suggests a practical application in VSNs : when the energy in the battery is insufficient to support all of the image processing in the VSN, the rate allocation from the previous time may be used as an suboptimal solution.

Table II
OPTIMAL RATE ALLOCATION USING SSIM

	τ_0			$\tau_0 + \Delta\tau$		
$C[t]$	$\mathbf{R}_{\tau_0}$	order	SSIM	$\mathbf{R}_{\tau_0}$	order	SSIM
5	(2,1,1,1)	1 / 4	0.469	(2,1,1,1)	1 / 4	0.489
6	(3,1,1,1)	1 / 10	0.587	(3,1,1,1)	1 / 10	0.615
7	(4,1,1,1)	1 / 20	0.647	(4,1,1,1)	1 / 20	0.699
8	(4,2,1,1)	2 / 35	0.693	(4,2,1,1)	2 / 35	0.740
9	(5,2,1,1)	2 / 56	0.731	(5,2,1,1)	2 / 56	0.791
10	(5,3,1,1)	5 / 84	0.772	(5,3,1,1)	5 / 84	0.830
11	(5,3,2,1)	12 / 120	0.802	(6,3,1,1)	5 / 120	0.856
12	(6,3,2,1)	11 / 161	0.830	(6,4,1,1)	10 / 161	0.878
13	(6,4,2,1)	18 / 204	0.856	(6,4,2,1)	18 / 204	0.898
14	(6,4,3,1)	29 / 246	0.878	(6,4,3,1)	29 / 246	0.915
15	(6,4,3,2)	44 / 284	0.893	(7,4,3,1)	19 / 284	0.929

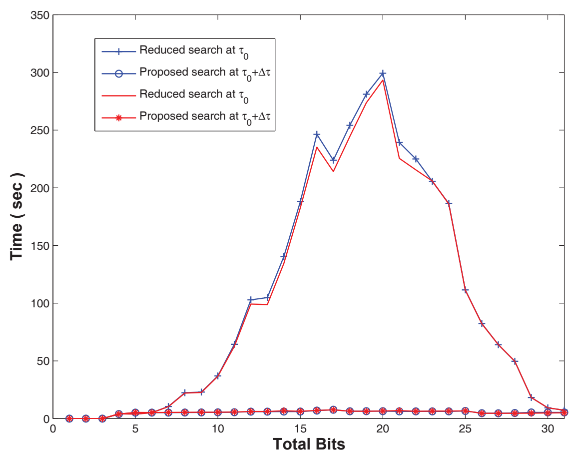


Figure 2. Computation time for each algorithm

In order to show the superiority of our algorithm, we have experimented with another scheme : reduced full search, which searches the bit combinations from order 1 to the order allocated equally to each bit. For example, when $C[t] = 12$, the reduced full search algorithm searches from (9,1,1,1) to (3,3,3,3). Fig. 2 shows the time spent for the proposed algorithm and the reduced full search. In the reduced full search, the required time increases until $C[t] = 20$, because the number of bit combinations satisfying the constraints in (5) is maximized at $C[t] = 20$. On the other hand, the proposed search spends almost the same amount of time. This result shows that our proposed search requires much less time than the reduced full search.

V. CONCLUSION

We described a DCT-based bit allocation scheme to get maximal image quality over a VSN, consisting of two features, "SLO (Searching from the Lowest Order)" and "SPP (Searching the Present optimal bit allocation from the Previous optimal bit allocation)". In the simulations, we showed the superiority of the proposed algorithm over the conventional approach, the reduced full search. In addition, since the optimal rate allocation does not frequently change with the image captured, energy and time can be saved by using the optimal rate allocation scheme conducted at the previous time.

VI. ACKNOWLEDGMENT

This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2010-(C1090-1011-0006))

REFERENCES

[1] I.F.Akyildiz and T.Melodia and K.R.Chowdhury, "A survey on wireless multimedia sensor networks," *Computer Networks*, vol. 51, no. 4, pp. 921-960, April 2007.

[2] P. Kulkarni, D. Ganesan, P. Shenoy, Q. Lu, "SensEye: a multi-tier camera sensor network," In *Proc. of ACM Multimedia*, pp. 229-238, Nov. 2005.

[3] Z.Wang, A.C.Bovik, H.R.Sheikh and E.P.Simoncelli, "Image Quality Assessment : From error visibility to structural similarity", *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004.

[4] S.S.Channappayya, A.C.Bovik and R.W. Heath, "Rate Bounds on SSIM Index of Quantized Images," *IEEE Tran. Image Processing*, vol. 17, no. 9, pp. 1624-1639, Sep. 2008.

[5] R. Puri, A. Majumdar, P. Ishwar, and K. Ramchandran, "Distributed video coding in broadband wireless sensor networks", *IEEE Signal Processing Mag.*, vol. 23, no. 4, pp. 94-105, April 2006.

[6] J. C. Dagher, M. W. Marcellin, and M. A. Neifeld, "A method for coordinating the distributed transmission of imagery", *IEEE Trans. Image Processing*, pp. 1705-1717, July 2006.