Distortion Management of Real-time MPEG-4 FGS Video over Downlink Multicode CDMA Networks

Guan-Ming Su, Zhu Han, Andres Kwasinski, Min Wu, K.J.Ray Liu, and Nariman Farvardin
Department of Electrical and Computer Engineering, University of Maryland, College Park.

Abstract— **In this paper, a protocol is designed to manage source rate/channel coding rate adaptation, code allocation, and power control to transmit real-time MPEG-4 FGS video over downlink multicode CDMA networks. We develop a fast adaptive scheme of distortion management to reduce the overall distortion received by all users subject to the limited number of codes and maximal transmitted power. Compared with a modified greedy method in literature, our proposed algorithm can reduce the overall system's distortion by at least 45%.**

I. Introduction

Transmitting real-time compressed video over CDMA networks is an emerging service. Compressed video exhibits a highly bursty rate due to the various complexities of different video contents and intra/inter coding mode. Many recent research works have concentrated on different aspects of this issue. A variable bandwidth retransmission scheme in an MC-CDMA system was proposed in [1]. Deep and Feng in [2] proposed a channel allocation policy by dynamically assigning more codes to an I frame in a multi-user MC-CDMA system. A joint rate and power allocation scheme for 3D-ESCOT scalable video codec was studied in [3]. An overview of current and future video over wireless was presented in [4]. The performance of CBR H.263 video over Nakagami fading channels in IS-95 CDMA systems for single-cell and multi-cell environment was studied in [5]. Chan *et. al.* [6] analyzed the capacity of a CDMA system supporting homogeneous H.263 video traffic. A multirate DS-CDMA system supporting heterogeneous services for QoS balance was studied in [7]. A scheme minimizing the overall power consumption of source/channel coding and transmission power was proposed in [8]. To provide all subscribers with satisfactory received qualities, we face a critical issue: the system's resources are limited. How to jointly perform the rate adaptation, code allocation, and power control to achieve required perceptual qualities of received video via distortion management becomes an important research problem.

In this paper, we study the resource allocation problem transmitting real-time MPEG-4 FGS video sequences over downlink multicode CDMA systems. We first design a protocol for the video transmission system over wireless. Then the distortion management is formulated as an optimization problem to achieve minimal overall distortion received by all users subject to the available number of codes and maximal power for transmission. We develop a fast distortion management algorithm to allocate resources to each user. Simulations show that the proposed algorithm reduces the distortion by at least 45%, compared with the modified greedy algorithm [9].

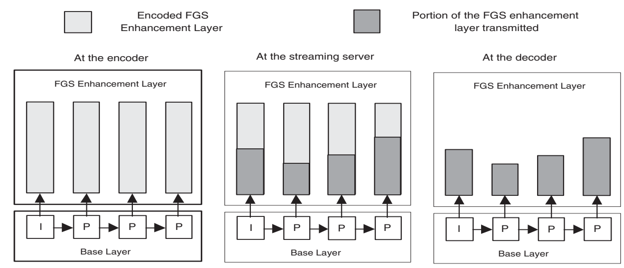

Fig. 1. Streaming Video using MPEG-4 FGS

This paper is organized as follows. The designed protocol is presented in Section II. In Section III, we have the system model and formulate the problem. In Section IV, we develop the resource allocation algorithm. Simulations are presented in Section V and conclusions are drawn in Section VI.

II. Protocol Description

In this section, we first briefly review the MPEG-4 FGS coding system. Then, we construct the protocol to implement video transmission over downlink multicode CDMA.

MPEG-4 Fine Granularity Scalability (FGS) coding [10] and Fine Granular Scalability Temporal (FGST) coding [11] are the new two-layer video techniques for delivering streaming video. FGS coding enables a video sequence to be encoded once, and transmitted/decoded at different rates according to the available bandwidth. The encoder generates a base layer at a low bit rate using a large quantization step and computes the residues between the original frame and the base layer. The bit planes of DCT transformed coefficients of these residues are encoded sequentially to form the FGS enhancement layer. The decoder can decode any truncated segment of the bitstream of FGS layer corresponding to each frame. The more bits the decoder receives and decodes, the higher the video quality is. Figure 1 illustrates a streaming video system using a FGS codec. The encoder encodes all bit planes for each video frame and lets the video server determine how many bits to send for each frame according to the channel condition. The decoder can decode the received truncated bitstream.

Figure 2 shows a block diagram of our proposed distortion management protocol to transmit FGS video over multicode CDMA. The protocol is implemented at the base station. The system resources, such as the number of codes and power, are managed to reduce the overall distortion. All users have their own FGS encoders to encode different real-time video programs. Those FGS encoders send the

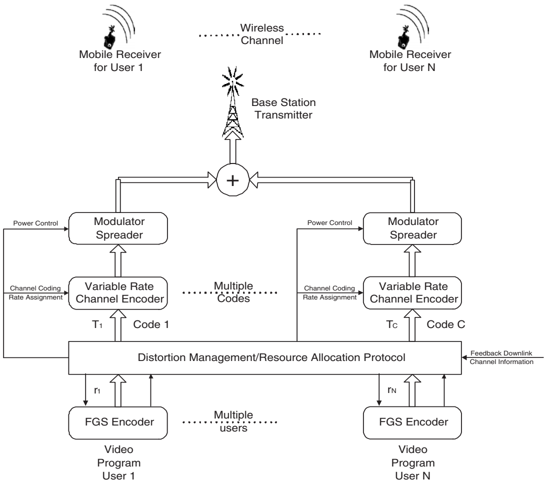

Fig. 2. Block Diagram for the Proposed Protocol

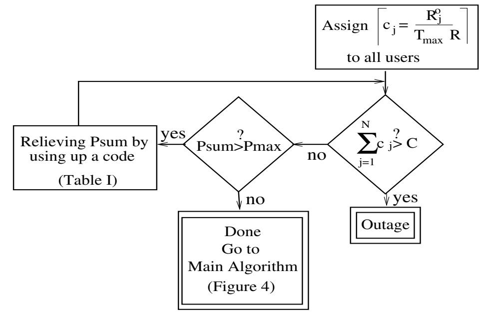

Fig. 3. Initialization Algorithm for Base Layer

rate-distortion (R-D) information to the proposed protocol. The protocol assigns a variable number of codes to each user according to his/her resource needs and channel conditions. For example, an I frame requires more codes than a P frame. Also, according to the feedback of downlink channel estimations, the protocol assigns the channel coding rates and power allocations to each code. In allocating resources, our goal is to maintain good video qualities, even when transmitting through a noisy channel with interference. Channel-induced errors affect qualities in an unpredictable way: for the same channel conditions, random errors may affect the received qualities in very different ways for different users. To avoid the uncertainty and maintain controllable video qualities, we use adaptive channel coding and power control to achieve a sufficiently small Bit Error Rate (BER). Because the number of codes and the overall transmitted power are limited, the challenge for the proposed protocol is how to efficiently allocate these resources such that the overall system distortion can be minimized.

III. System Model and Problem Formulation

Consider a single cell MC-CDMA system with N users and a total of C codes. We assume the system is synchronous and each user is assigned a set of unique pseudo-random codes. Because of the multipath effect [12], the orthogonality among codes may not be guaranteed. Consequently, each mobile user is subject to the interference from other users in the cell. If the i^{th} code is assigned to user j, the received SINR is:

$$\Gamma_i = \frac{W}{R} \frac{P_i G_j}{G_j \sum_{k=1, k\neq i}^{C} \alpha_{ki} P_k + \sigma^2} \tag{1}$$

where W is the total bandwidth and is fixed, R the transmit rate, P_i the transmitted power from the base station for code i, α_{ki} orthogonality factor between codes, G_j the j^{th} user's path loss, and σ^2 the thermal noise level that is assumed to be the same at all mobile receivers. The ratio W/R is the processing gain.

Our goal is to maintain the BER after channel decoding below a threshold such that the video quality is controllable by the MPEG-4 FGS video compression with error resilient and concealment techniques. To reach this goal, the received SINR should be no less than a targeted SINR, which is a function of channel coding rate. Note that for a fixed transmit rate, a reduction in source encoding rate allows for inserting more channel protection. Through simulations using Rate Compatible Punctured Convolutional (RCPC) codes [13], we have found that for our application the targeted SINR can be accurately approximated as a function of channel coding rate by

$$\gamma_i = 2^{AT_i + B} \tag{2}$$

where γ_i is the required targeted SINR. A and B are the parameters of the error control coding, and T_i is the channel coding rate with range $[T_{min}, T_{max}]$.

Existing video schemes using a single-layer video codec often explicitly employ R-D models and an exponential or a polynomial R-D model is frequently used [14], [15]. The MPEG-4 FGS codec is a two-layer scheme and its enhancement layer is encoded bit plane by bit plane. For a given bit plane in a frame, if the video is spatially stationary so that the length of the entropy encoded FGS symbols in all blocks is similar to each other, the decoded bit rate and the corresponding amount of reduced distortion will have an approximately linear relationship over the bit rate range of this bit plane. Previous studies in [15], [16] show that a piecewise linear function is a good approximation to the R-D curve of FGS video at the frame level. This piecewise linear function model can be summarized as:

$$D_{n,j}(r_{n,j}) = M_{n,j}^k (r_{n,j} - R_{n,j}^k) + E_{n,j}^k,\ k = 0, \ldots, p-1, \tag{3}$$

$$\text{with } M_{n,j}^k = \frac{E_{n,j}^{k+1} - E_{n,j}^k}{R_{n,j}^{k+1} - R_{n,j}^k}, R_{n,j}^k \leq r_{n,j} \leq R_{n,j}^{k+1},$$

where p is the total number of bit planes, and $E_{n,j}^k$, $R_{n,j}^k$, and $r_{n,j}$ denote the distortion measured in mean square error (MSE) after completely decoding the first k DCT bit planes, the corresponding bit rate, and the overall decoded bit rate for the j^{th} user's distortion of the n^{th} frame, respectively. Note that $E_{n,j}^0$ and $R_{n,j}^0$ represent the distortion and source rate of the base layer, respectively. Because DCT is a unitary transform, all $(R_{n,j}^k, E_{n,j}^k)$ R-D pairs can be obtained during the encoding process without decoding the compressed bitstreams. Since we allocate resources frame by frame, we omit n from the notation for simplicity.

In this multicode system, we denote a_{ij} as an indicator to specify whether the i^{th} code is assigned to user j. The maximal overall power is P_{max}, and each user's throughput, r_j, should be larger than the base layer rate R_j^0 to guarantee the baseline quality and smaller than the maximum source rate R_j^P. We formulate this optimization problem as:

$$\min_{T_i, a_{ij}} \sum_{j=1}^{N} D_j \tag{4}$$

$$\text{subject to} \begin{cases} \sum_{j=1}^{N} a_{ij} \leq 1, a_{ij} \in \{0,1\}, \forall i; \\ P_{sum} = \sum_{i=1}^{C} P_i \leq P_{max}; \\ R_j^0 \leq r_j = R \sum_{i=1}^{C} a_{ij} T_i \leq R_j^P, \forall j. \end{cases}$$

The difficulty to solve (4) lies in the power constraint and code constraint. Each user can reduce his/her distortion by increasing transmitted power. Because different users experience different channel conditions, users will need different increases of power to have the same distortion reduction. However, the overall power at the base station is limited. We need to allocate power to each user efficiently. Moreover, each user may transmit an I frame or a P frame and may encounter different complexity of the video content at each time instance. Allocating a code to different users results in different reductions of the overall distortion. In the next section, we will develop algorithms to efficiently allocate the limited power and codes to reduce the overall distortion.

IV. Distortion Management Algorithm

There are two main stages in our proposed algorithm. At the first stage, we allocate the resources for delivering the base layer data to provide the baseline video quality for each user. Some FGS layer data is then delivered to reduce the overall distortion by adjusting system's resource allocation at the second stage. The goal for the adjustment is to fully utilize the code and power resources and to avoid exhausting only one resource first while having the other resource left, which leads to local optima. The proposed algorithm can overcome the problem mentioned above by keeping a balance between code and power allocation during the process of resource allocation.

A. Base Layer

Figure 3 shows the initialization procedure, where codes, channel coding rates, and power are assigned to each user so that all users can transmit the base layer rate R_j^0 only, while the power constraint is satisfied. First, we use the maximal channel coding rate T_{max} and assign the number of codes c_j for each user, such that all the base layers can be transmitted. If there is no code left, an outage will be reported, indicating that there are too many users in the system and there are no resources even for accommodating the base layers only. If there are enough codes for the base layer, we will judge whether the power constraint is violated. If no, the initialization is done and we go to the main algorithm for FGS layer. Otherwise, we will relieve P_{sum} by assigning a code once a time to the user who can reduce the power most while the distortion is fixed, until the power constraint is satisfied.

Table I: P_{sum} Relieve Algorithm

1. For hypothesis $j = 1$ to N:
   - Assign a candidate code to user j.
   - Calculate the optimal T_i for all codes assigned to user j including the candidate code, such that P_{sum} is reduced most, while r_j is unchanged.
2. Pick the user with the largest reduced P_{sum} and assign him/her a real code.

We can approximate P_{sum} as follows. Depending on which user the code is assigned to, we define

$$Y_i = \begin{cases} 0, & \text{if code is not assigned;} \\ \frac{2^{AT_i+B}R}{W} = \frac{P_i G_j}{G_j \sum_{k \neq i} \alpha_{ki} P_k + \sigma^2}, & \text{for user j.} \end{cases} \tag{5}$$

Since the processing gain W/R is large and Y_i is small, P_{sum} can be approximated as:

$$\begin{aligned} P_{sum} &= \mathbf{1}^T[\mathbf{I}-\mathbf{F}]^{-1}\mathbf{u} \approx \mathbf{1}^T[\mathbf{I}+\mathbf{F}]\mathbf{u} \\ &= \sum_{i=1}^{C} \frac{\sigma^2 Y_i}{G_i} + \sum_{i=1}^{C}\sum_{k\neq i}^{C} \frac{\sigma^2 \alpha_{ki} Y_i Y_k}{G_k}, \end{aligned} \tag{6}$$

where $\mathbf{1} = [1 \ldots 1]^T$, $\mathbf{u} = [u_1, \ldots u_C]^T$ with $u_i = \sigma^2 Y_i / G_i$, and $[\mathbf{F}]_{ij} = 0$ if $j = i$; $[\mathbf{F}]_{ij} = \alpha_{ji} Y_i$ if $j \neq i$.

The P_{sum} relieve algorithm is shown in Table I. Before we assign a real code to a specific user, we make N hypotheses. For the j^{th} hypothesis, we assign a candidate code to the j^{th} user and the settings of the other users keep unchanged. The j^{th} user will keep the amount of his/her source coding rate, r_j, unchanged, but redistribute r_j to his/her existing assigned codes with the candidate code. Consequently, the channel coding rates for those codes are reduced, so that the required SINR and the overall transmitted power are reduced. The problem can be formulated as an optimization problem to minimize P_{sum} subject to a fixed distortion by allocating the source coding rates and channel coding rates to the already assigned codes plus the candidate code. This problem can be solved using water filling method. First, we assign the candidate code with T_{max}. So the current throughput is larger than r_j. Since Y_i is a monotonic increasing function of T_i, we can reduce P_{sum} by searching the code with the largest $|g_i^T = \frac{\partial P_{sum}}{\partial T_i}|$ and reduce the channel coding rate of this code. The algorithm repeats the searching procedure until the throughput is equal to r_j. From all hypotheses, the user reducing P_{sum} by the highest amount is selected and assigned a real code.

B. FGS Layer

After initialization, we apply the distortion management algorithm in Figure 4 to efficiently allocate resources for FGS layer to reduce users' overall distortion. The algorithm first decides whether all codes are used up. If yes, we use the remaining power to reduce the distortion as described in Table III. Otherwise, different sub-algorithms are applied by assigning a new code to reduce the power or distortion depending on whether the system is power unbalanced. The whole distortion management algorithm for FGS layer is terminated under two conditions: First, all available codes are distributed and the power is within

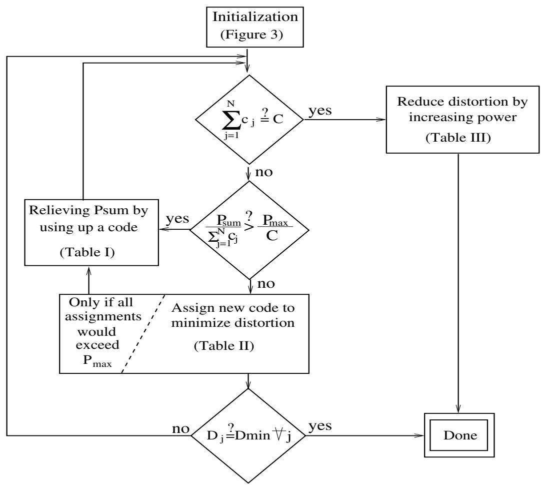

Fig. 4. Distortion Management Algorithm for FGS Layer

the range. This is the situation when the system is heavy loaded. Second, all the users have the minimal distortion. This is the light-load situation.

The criterion for judging whether the system is power unbalanced is the average consumed power per assigned code. If the current P_{sum} over the assigned code $\sum_{j=1}^{N} c_j$ is greater than P_{max} over the maximal number of codes C, the system is power unbalanced. Then, the algorithm in Table I is applied to reduce the power. Otherwise, we apply the algorithm listed in Table II to reduce the distortion.

In Table II, we reduce the overall distortion by assigning one code to a user at a time. Before we assign a real code to a specific user, we make N hypotheses. For the j^{th} hypothesis, we assign a candidate code with the channel coding rate T_{max} to the j^{th} user and the settings of the other users keep unchanged. Then, we calculate the overall required power for transmission, P_{sum}, using (6), and the reduced distortion of the received video as follows:

$$\Delta D_j = D_j(r_j) - D_j(r_j + RT_{max}) \qquad (7)$$

where r_j is the user's current source rate. If P_{sum} is smaller than P_{max}, we add this hypothesis into the candidate list. Then among all candidate hypotheses, we assign a real code to the user who can reduce the distortion by the highest amount. If there is no candidate in the list, it means the overall distortion cannot be further reduced since the required power exceeds P_{max}. To facilitate further distortion reduction in the successive iterations, the power needs to be reduced using the algorithm in Table I.

When all the codes are assigned and there is some transmitted power left, we can further reduce the distortion by increasing the power. The algorithm is listed in Table III. We make C hypotheses. For hypothesis i, we check whether the channel coding rate of code i, T_i, is less than T_{max}. If no, we check the next hypothesis. Otherwise, we increase T_i by a discrete step ΔT_i and keep the settings of the rest C-1 codes unchanged. If this code belongs to the j^{th} user, we calculate the reduced distortion ΔD_j and increased overall power ΔP_{sum}. We add this hypothesis in the candidate list. Among these candidates, our strategy is to pick the

Table II: Code Assignment to Reduce Distortion

1. For hypothesis $j = 1$ to N:
• Assign user j a candidate code, analyze ΔD_j,P_{sum}
• If $P_{sum} < P_{max}$, add hypothesis j to candidate list.
2. If there is no candidate user, do not assign the code and go to the algorithm in Table I.
3. Among the candidates, choose the one with the largest ΔD_j and assign a real code to user j.

Table III: Distortion Reduction by Increasing Power

1. For hypothesis $i = 1$ to C:
• If T_i of code i is equal to T_{max}, do next hypothesis.
• For code i, calculate the corresponding decrease in channel coding rate of one discrete step, ΔT_i.
• Given ΔT_i, calculate Δr_j, ΔD_j, and ΔP_{sum}. If $P_{sum} < P_{max}$, add hypothesis i to candidate list.
2. If no candidate left, exit. Otherwise, choose the code with the largest $\lvert\Delta D_j/\Delta P_{sum}\rvert$ and change the channel coding rate to the chosen code.
3. Empty candidate list. Go to step 1.

code with the largest $|\Delta D_j/\Delta P_{sum}|$ and set $T_i = T_i + \Delta T_i$. The above process is repeated until there is no power left.

V. Simulation Results

The simulations are set up as follows. The bandwidth is 7.5 MHz. The spreading factor is 64. The path loss factor is 4. The delay profile is typical urban, and we use the average orthogonality factor in [12]. The noise power is 10^{-9} Watts, and the maximal transmitted power is 280 Watts. The mobile is uniformly distributed within the cell with radius from 20 m to 700 m. We use RCPC codes with a memory 4, puncturing period 8, and mother code rate 1/4 [13]. Our experimental results show that to achieve BER= 10^{-6} using the FGS codec (including base and FGS layer), parameters (A, B) in (2) are (4.4, -1.4). The video refresh rate is 15 frames per second. We concatenate 15 classic video sequences (*Akiyo*, *Carphone*, *Claire*, *Coastguard*, *Container*, *Foreman*, *Grandmother*, *Hall objects*, *Miss American*, *Mother and daughter*, *MPEG4 news*, *Salesman*, *Silent*, *Suzie*, and *Trevor*) with temporal down sampling factor 2 to form a basic testing video sequence source of 2775 frames. The base layer is generated by MPEG-4 encoder with a fixed quantization step of 30 and the GOP pattern is 14 P frames after one I frame. All frames of FGS layer have up to six bit planes. The content program of the i^{th} user is 100 frames long and starts from frame 173×(i-1)+1 of the concatenated video source.

Figure 5 shows a simulation result for the convergence track of the overall power and distortion with the number of assigned codes by using the proposed algorithm. After initialization (shown at A), the base layer of each user is allocated and 17 codes are assigned. The overall distortion (shown at A') is large because only the base layer is transmitted. While the system is power unbalanced, i.e. the operating point (such as position A) is above the balanced resource allocation line, we apply the power relieve algorithm in Table I to reduce the power while keeping the distortion fixed. When the system is not power unbalanced

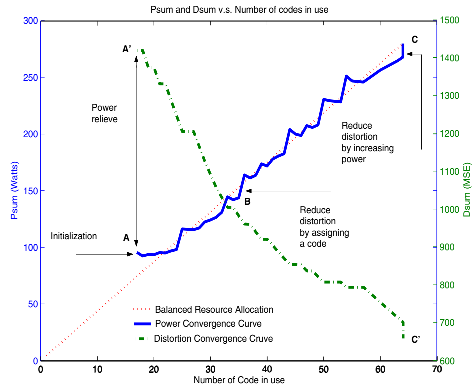

Fig. 5. Power and Distortion vs. the number of Assigned Codes

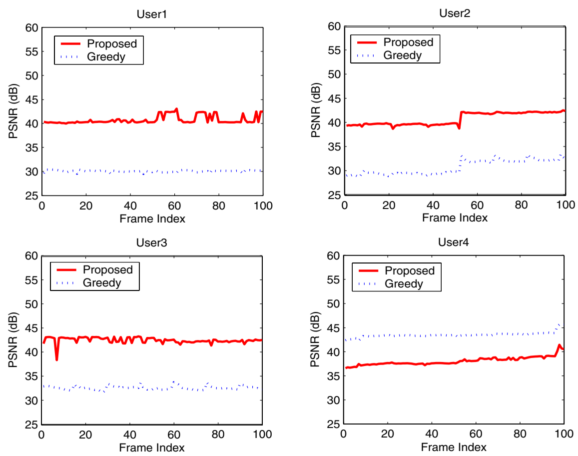

Fig. 6. PSNR Results for User 1 to User 4 vs. Frame Number

(such as position B), we assign codes to reduce distortion (consequently, the required power is increased) using the algorithm in Table II, until all the codes are used up. Finally, we use the algorithm in Table III to further reduce distortion by the remaining power quota (shown at C).

We compare the proposed algorithm with a modified greedy approach [9]. This modified approach is similar to our proposed framework, but uses a greedy approach for the code assignment in FGS layer. For each iteration, this greedy algorithm tries to assign a candidate code to every user, calculates $|\Delta D_j/\Delta P_{sum}|$, and assigns a new code to the user with the largest value. Figure 6 shows the PSNR results in a four-user system. The first three users receive better video qualities using the proposed algorithm. Since the greedy algorithm assigns codes to the users who can use the least power to obtain the largest decreased distortion, which is the fourth user in this example, it cannot reduce the overall distortion much. Figure 7 shows the number of users v.s. average of the total distortion D_{sum} ($D_{sum} = \sum_{j=1}^{N} D_j$) over 100 frames from 50 different mobiles' locations. The simulation results demonstrate that the average D_{sum} of the proposed algorithm outperforms that of the greedy algorithm by at least 45%.

VI. Conclusions

In this paper, we have developed a protocol to transmit real-time MPEG-4 FGS video over downlink multi-

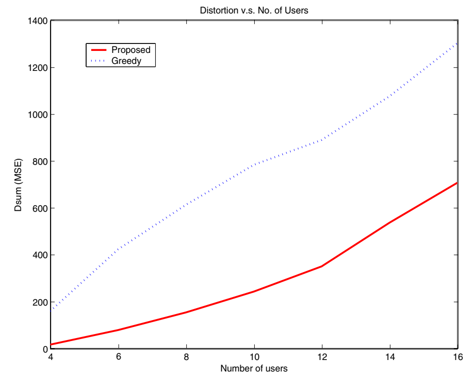

Fig. 7. Performance Comparison of the proposed and greedy scheme

code CDMA networks. The resource allocation is formulated as an optimization problem to minimize the overall received distortion of all users subject to the power and the number of codes constraints. To fully utilize the limited resources, we propose a fast distortion management algorithm to jointly allocate source rates, channel coding rates, codes, and power. Experimental results show that our proposed approach provides an efficient solution for sending video over CDMA system.

References

[1] M.R. Hueda, et. al, "Enhanced-performance video transmission in multicode CDMA wireless systems using a feedback error control scheme", *IEEE Globecom 2001*, pp.619 -626, Nov. 2001.

[2] B. Deep and W.-C. Feng "Adaptive code allocation in multicode-CDMA for transmitting H.263 video", *IEEE WCNC*, 1999.

[3] S. Zhao, Z. Xiong, and X. Wang, "Optimal resource allocation for wireless video over CDMA networks", *ICME* 2003.

[4] L. Hanzo, et. al, "Wireless video communications, second to third generation systems and beyond", *IEEE Press*, 2001.

[5] N.H.L. Chan and P.T. Mathiopoulos, "Efficient video transmission over correlated Nakagami fading channels for IS-95 CDMA systems", *IEEE JSAC*, vol.18, no.6, pp.996-1011, Jun. 2000.

[6] Y. S. Chan and J.W. Modestino, "A joint source coding-power control approach for video transmission over CDMA networks", *IEEE JSAC.*, vol.21, no.10, pp.1516-1525, Dec. 2003

[7] L. Wang , et. al., "Optimal power allocation for a multi-rate packet CDMA system with multi-media traffic", *IEEE WCNC99*.

[8] Q. Zhang, Z. Ji, W. Zhu, and Y.-Q. Zhang, "Power-minimized bit allocation for video communication over wireless channels", *IEEE Trans. on Circ. and Sys. for Video Techn.* , Jun. 2002.

[9] Z. Han, A. Kwasinski, K.J.R. Liu, and N. Farvardin, "Pizza party algorithm for real time distortion management in downlink single-cell CDMA systems", *Allerton Conference*, 2003.

[10] H.M. Radha, M. van der Schaar, and Y. Chen, "The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP", *IEEE Transactions on Multimedia.* , Mar 2001.

[11] M. van der Schaar and H.M. Radha, "A hybrid temporal-SNR fine-granular scalability for internet video", *IEEE Trans. on Circuits and Systems for Video Technology.* , Mar. 2001.

[12] O. Awoniyi, N.B. Mehta, and L.J. Greenstein, "Characterizing the ortogonality factor in WCDMA downlinks", *IEEE Trans. on Wireless Communications*, vol.2, no.4, pp.621-625, Jul. 2003.

[13] J. Hagenauer, "Rate compatible punctured convolutional (RCPC) codes and their applications", *IEEE Transactions on Communications*, vol. 36, pp.389-399, Apr. 1988.

[14] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications", *IEEE Trans. on Circuits and Systems for Video Technology* , pp. 172-185, Feb. 1999.

[15] X.M. Zhang, A. Vetro, Y.Q. Shi, and H. Sun, "Constant quality constrained rate allocation for FGS-coded videos", *IEEE CSVT*, pp. 121-130, Feb. 2003.

[16] L. Zhao, J. Kim and C.-C. J. Kuo, "MPEG-4 FGS video streaming with constant-quality rate control and differentiated forwarding", *SPIE, Conf. on Visual Comm. and Image Proc., 2002.*