Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Optimizing inter-view prediction structures for multi-view video coding using simulated annealing*

Zheng ZHU†1, Dong-xiao LI[1,2], Ming ZHANG[1,2]

([1]*Institute of Information and Communication Engineering, Zhejiang University, Hangzhou 310027, China*)

([2]*Zhejiang Provincial Key Laboratory of Information Network Technology, Hangzhou 310027, China*)

†E-mail: PAL106@163.com

Received Jan. 16, 2010; Revision accepted Mar. 23, 2010; Crosschecked Dec. 30, 2010

Abstract: New video applications, such as 3D video and free viewpoint video, require efficient compression of multi-view video. In addition to temporal redundancy, exploiting the inter-view redundancy is crucial to improve the performance of multi-view video coding. In this paper, we present a novel method to construct the optimal inter-view prediction structure for multi-view video coding using simulated annealing. In the proposed model, the design of the prediction structure is converted to the arrangement of coding order. Then, a simulated annealing algorithm is employed to minimize the total cost for obtaining the best coding order. This method is applicable to arbitrary irregular camera arrangements. As experiment results reveal, the annealing process converges to satisfactory results rapidly and the generated optimal prediction structure outperforms the reference prediction structure of the joint multi-view video model (JMVM) by 0.1–0.8 dB PSNR gains.

Key words: Multi-view video coding, Prediction structure, Simulated annealing
doi:10.1631/jzus.C1000016 **Document code:** A **CLC number:** TN941.3

1 Introduction

Recently, the convergence of new technologies from computer graphics, computer vision, multimedia, and related fields also enabled the development of new types of media, like 3D video, which expands the user's sensation far beyond what is offered by traditional media (Smolic *et al*., 2006). More and more technologies have been launched to support 3D video systems. Among them, the multi-view video coding (MVC) is the key technology for various applications, including free-viewpoint television, 3D television, and immersive teleconferencing (Kalva *et al*., 2006). Unlike traditional 2D video, multi-view video is generated by multiple cameras simultaneously capturing the same scene from different viewpoints. Since this approach creates large amounts of data to be stored or transmitted to the user, efficient compression techniques are essential for realizing these applications (Merkle *et al*., 2007).

Contrary to the single view video, a multi-view video contains not only temporal redundancy, but also large amounts of inter-view redundancy. Efficient exploitation of temporal and inter-view correlations is critical to MVC, which relies on the design of prediction structure. To date, many prediction structures for MVC have been proposed, aiming at overall compression efficiency improvement. Among them, the 'hierarchical B pictures' proposed by Heinrich Hertz Institute (HHI) have been selected by the Joint Video Team (JVT), for its excellent performance, as the reference view-temporal picture structure of the joint multi-view video model (JMVM) (Muller *et al*., 2006). Most of these prediction structures, however, were manually designed for simple and static camera arrangements, especially for equispaced 1D camera arrays. To achieve better compression efficiency and applicability in arbitrary irregular camera arrange-

* Project supported by the National Natural Science Foundation of China (No. 60802013) and the Zhejiang Provincial Natural Science Foundation of China (No. Y106574)

ments, a method of minimum spanning tree was proposed by Li *et al*. (2007) and Kang *et al*. (2007). By representing each picture as a node and the prediction costs as the weights of edges, they adopted the graph theory to find the optimal prediction structure. B pictures cannot, however, be involved in their prediction structures because they would yield loops. In general, B pictures are more efficient than I and P pictures in compression performance, and should be considered in prediction structure construction.

In this paper, a novel method is proposed to construct the optimal prediction structure to include B pictures. A simulated annealing (SA) algorithm is employed in the model to search for the best coding order of views. SA is an optimization process derived from the physical process of cooling molten material down to the solid state (Kirkpatrick *et al*., 1983). It searches for the minimum energy state of the cost function regardless of the shape of the function and can escape from local minima with hill-climbing (Bohachevsky *et al*., 1986). Because of these features, SA has been widely used for various combinatorial and other optimization problems (Laarhoven and Aarts, 1988). SA was applied in the proposed model to investigate the following problems: (1) Given the prediction costs, what is the best way to determine the optimal inter-view prediction structure that includes B pictures? (2) How much improvement can be achieved from inter-view prediction over JMVM?

2 Prediction structure for multi-view video coding

The main consideration regarding efficiency of MVC is the efficiency gain of inter-view/temporal prediction. Statistical analysis over a large set of multi-view sequences shows that temporal prediction is always the most efficient prediction mode and interview is more efficient than inter-view/temporal mixed modes at large (Muller *et al*., 2006). Hence, it is reasonable to adopt an 'orthogonal' view-temporal framework, wherein a picture can be used to predict the following coding pictures at the same time-point or in the same view. The reference prediction structure of MVC is illustrated in Fig. 1 for a sequence with eight cameras and a group of pictures (GOP) length of eight (Merkle *et al*., 2007).

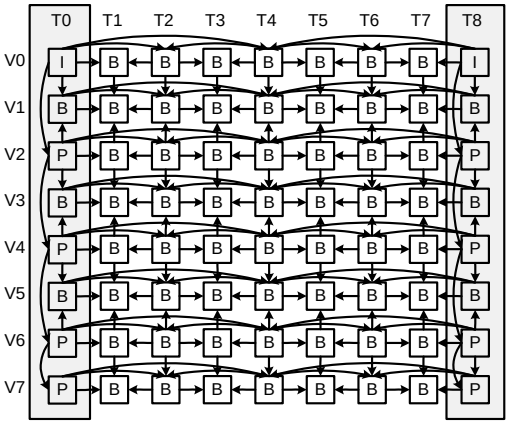

Fig. 1 Reference prediction structure for multi-view video coding (MVC) with eight cameras and a group of pictures (GOP) length of eight

V*n* denotes the individual view sequence and T*n* the consecutive time-point

A prediction structure of the 'orthogonal' framework can be separated into prediction structures in temporal and inter-view dimensions. In temporal dimension, the proposed prediction structure uses hierarchical B pictures for each view as the reference prediction structure. Hierarchical B pictures provide significantly improved rate-distortion (RD) performance when the quantization parameters for the various pictures are assigned appropriately (Schwarz *et al*., 2006). To achieve good compression performance, the reference prediction structure uses the correlation between the views by applying inter-view prediction to every second view. This design is effective for the simple and static 1D camera arrangements, especially for the case when only the adjacent views are the suitable references.

This study aims to exploit the inter-view correlation to the best of its potential for applications in arbitrary irregular camera arrangements. It is very complicated, however, to search for the optimal prediction structure, such that optimization methods are difficult to directly apply. In this case, we determine the optimal coding order of views rather than directly search the prediction structure. Given the coding costs, one prediction structure with minimum cost corresponding to the coding order can be built. Fig. 2 illustrates an arbitrary coding order for an example of eight views. For the first coding view V3, coding mode of I is the only choice. For V5, besides the I view, it can be coded as a P view with V3 as its

reference. For V0, two reference views (V3 and V5) can be used, and thus the coding mode of B can be chosen. By parity of reasoning, the views with later coding order can take any previous view as reference. For simplification, only two-reference prediction is taken into account for B views here.

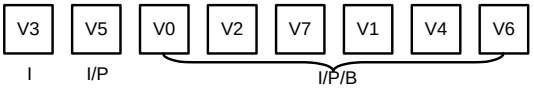

Fig. 2 Arbitrary coding order of an example of eight views

By choosing the best coding mode and references that can be used for each view, we can easily obtain the optimal prediction structure corresponding to each coding order. The optimization method is aimed to minimize the total cost to find the best coding order. Therefore, the optimal prediction structure corresponding to the best coding order can be obtained. We consider the scenario where coding costs are given, and adopt the output bits at a predefined distortion as the measure of coding costs. To reduce the computation complexity, other measures can be used instead. This work is concentrated on the optimization of the searching process. In the proposed model, SA is employed as a simple and effective method for finding these good results, which will be detailed in the next section.

3 Application of simulated annealing

3.1 Overview of simulated annealing

The procedure of SA employs methods originating from statistical mechanics to find global minima of a given bounded cost function with large degrees of freedom (Zomaya, 2001). The annealing process can start from any initial state in the domain of interest. According to the selected cost function, the energy of the current state, E_0, is calculated. Then a constraint-based new state is generated from the current one, with an energy of E_1. Let ΔE be the energy change of state, $\Delta E=E_1-E_0$. The next state is decided according to the Metropolis criterion (Metropolis *et al*., 1953). If the new state is better than the current one ($\Delta E\leq 0$), it is accepted unconditionally and becomes the next current state. Otherwise ($\Delta E>0$), the new state is not rejected outright but accepted with a certain probability. For example, if $\exp(-\Delta E/T)$ is greater than a random number uniformly distributed in (0, 1), the new state is accepted, where T is the control factor 'temperature'. This acceptance of a worse state helps the SA algorithm to escape from local minima. The temperature is high at the beginning to avoid local optima, and thus the probability of acceptance of a worse state remains high. At each T, a series of random new states are generated for selection. Then, an annealing scheme is applied by decreasing T according to some predefined schedule, which lowers the probability of acceptance of a worse state. After some point, a new state is no longer accepted unless it is better than the current one. Fig. 3 outlines the flow chart of the SA algorithm.

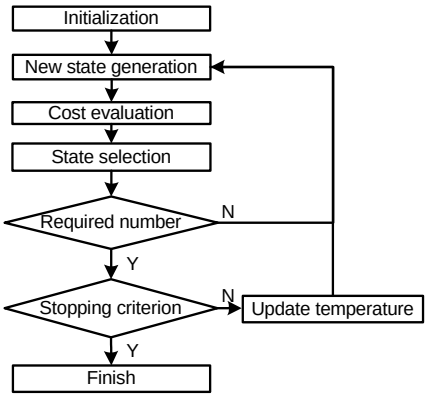

Fig. 3 Flow chart of the simulated annealing (SA) algorithm

3.2 Proposed model using simulated annealing

In this work, a novel model is presented to construct the optimal prediction structure. The SA algorithm is employed in the model to search for the best coding order of views. In the model, the permutation of views forms the domain of states. For an example of a view number of N, the number of different orders is as many as factorial of N. The initial state is set to be an arbitrary order of views, e.g., S_0={V3, V5, V0, V2, V7, V1, V4, V6}, where Vn (n=0, 1, …, N−1) denotes the individual view. Then, two views of the current set S_0 are randomly selected and their positions are exchanged with each other to generate the new state S_1, e.g., S_1={V3, V5, V0, V1, V7, V2, V4, V6}. Summing up the costs of all views yields the energy of a state, $E=\sum_{i=1}^{N} c_i$, where c_i denotes the

cost of the *i*th view of the state. Each view chooses the minimum cost from the coding modes and references, which can be used as its current cost. The next state is selected from S_0 and S_1 by applying ΔE to the Metropolis criterion, as described above.

For the convenience of investigation, we arrange the costs into two tables. The first table $\boldsymbol{C}_1$ is an $N\times N$ matrix, used for I and P coding costs. The element of the *i*th row and *j*th column denotes the prediction cost of view V*i* using view V*j* as the reference, and the elements in the diagonal denote the I coding costs. The second table $\boldsymbol{C}_2$ is used for B coding costs, with the size of $N\cdot \mathrm{C}_{N-1}^2$, where C_{N-1}^2 denotes the combination of two elements selected from $N-1$. Each element of $\boldsymbol{C}_2$ contains not only the value of cost but also the information of two references. The B coding costs of view V*i* are sorted ascendingly in the *i*th row of $\boldsymbol{C}_2$. A two-mode searching mechanism is designed for fast calculation. Besides sequential searching, $\boldsymbol{C}_2$ supports searching by index (Fig. 4). Sequential searching is an efficient mode for the views with later coding order, whereas indexical searching is better for the previous views. Take the state S_0 for example. V6 is the last view that can reference any other view; thus, the first element of the corresponding row is exactly the needed cost. As for V4, the only view that cannot be adopted as the reference is V6. Thus, the sequential searching is implemented until the first element is found that does not include V6 as a reference. For the third view V0, however, only two views (V3 and V5) can be used as a reference, so we interpret the reference views into the index in $\boldsymbol{C}_2$. For V2, three views can be used as a reference and the number of combinations of references is $\mathrm{C}_3^2=3$. The indexical searching will check three elements in $\boldsymbol{C}_2$ and find the minimum one as the cost for V2.

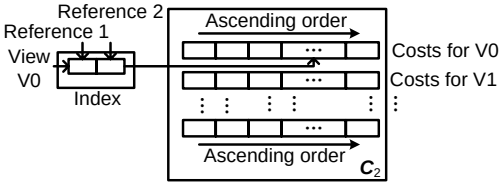

Fig. 4 Two-mode searching mechanism for B coding costs

To determine the critical number of views that adopt indexical searching, the following equation is used:

$$\mathrm{C}_{n-1}^2=\sum_{i=1}^{\mathrm{C}_{N-1}^2-(\mathrm{C}_{n-1}^2-1)}\mathrm{C}_{\mathrm{C}_{N-1}^2-i}^{\mathrm{C}_{n-1}^2-1}\cdot i\Big/\mathrm{C}_{\mathrm{C}_{N-1}^2}^{\mathrm{C}_{n-1}^2}, \tag{1}$$

where C_{n-1}^2 denotes the number of inquiring times by indexical searching for the cost of the *n*th view of the state, and the right hand side of Eq. (1) denotes the mathematical expectation of the number of inquiring times by sequential searching when the elements of $\boldsymbol{C}_2$ are assumed as following a uniform distribution. By solving this equation, the critical number n ($3\le n\le N$) is obtained. Note that n may be a decimal. The maximum integer that is smaller than or equal to n is adopted. For example, with N=8 views, the inquiring times of two searching modes are shown in Fig. 5. From the point of intersection, the critical number of four is derived. Thus, given a state with eight views, the costs for the first and second views are obtained from $\boldsymbol{C}_1$, and the costs for the third and fourth views are inquired in $\boldsymbol{C}_2$ by indexical searching. The costs for the remaining views are inquired in $\boldsymbol{C}_2$ by sequential searching.

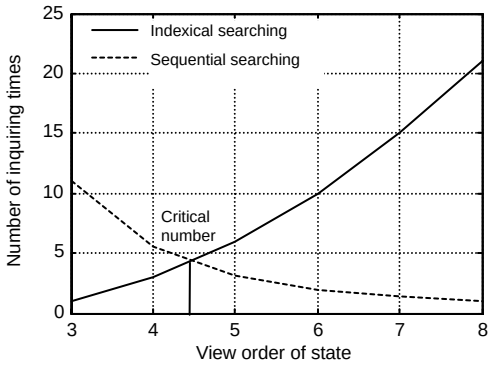

Fig. 5 Inquiring times of two searching modes for the case of eight views

In addition to the concise data structure presenting the problem and the well-formulated cost function, the design of an efficient annealing algorithm is influenced by the following factors: the initial temperature T_0, the rate at which the temperature is lowered, the number of iterations needed at any given temperature, and the stopping criterion of annealing. To determine the initial temperature, we use the ratio of the initial acceptance init_accept=exp($-\overline{\Delta E^+}/T_0$), where $\overline{\Delta E^+}$ denotes the increment of the energy averaged in init_iter times of random perturbations of

state. It is recommended that the value of init_accept be in the range 0.5<init_accept<0.9, because high probabilities usually cause inefficient evaluation and long annealing time while low probabilities make it difficult to escape from local minima. The temperature is controlled by a decreasing factor α: $T_{k+1}=\alpha \cdot T_k$. The number of iterations needed for each value of temperature is attached to the number of views: num_iter=$\beta \cdot N$. As for the stopping criterion, we use a tolerance factor ε to check the difference between the maximum energy E_{max} and minimum energy E_{min} of the accepted states at some temperature. When the difference between E_{max} and E_{min} satisfies $(E_{max}-E_{min})/E_{min}<\varepsilon$, the annealing stops.

4 Results

To evaluate the proposed method, an experiment platform was built on JMVM8.0 and eight publicly available test sequences for MVC common test conditions were used (Su *et al*., 2006). First, the coding costs based on the measure of output bits were generated by JMVM software. In this step, GOPSize and IntraPeriod were both set to 2, and the coding results of the anchor time points were used. To obtain relatively stable coding costs, 10 anchor time points were adopted for average. Then the SA algorithm was performed to search for the optimal inter-view prediction structures. For initial temperature decision, the ratio of initial acceptance, init_accept, was set to 0.8, and the number of initial random perturbation times, init_iter, was set to be num_iter. The decreasing factor α was set to 0.5 for controlling the temperature and β was set to 3 for deciding num_iter. The stop tolerance factor ε was set to 0.001. These factors were found empirically for rapid convergence. Fig. 6 illustrates the energy decay in the annealing process for sequence 'ballroom', where the dashed line denotes the global minimum energy obtained by the exhaustive search. It is obvious that the SA algorithm achieved fast convergence to the global minima.

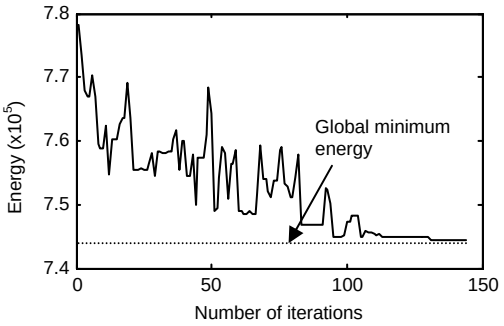

Fig. 6 Energy decay in the annealing process for sequence 'Ballroom'

Since SA is a stochastic optimization algorithm, it is possible that the result is different each time SA is implemented. Therefore, we applied SA to each test sequence for 10 times with the same parameter setting, and the best and worst results are listed in Table 1. For eight views, the SA algorithm converged to satisfactory results in less than 300 iterations most of the time (0.007:1 compared to the number of iterations in the exhaustive search). For 16 views, the number of iterations was less than 500 (10^{-11}:1 compared to the number of iterations in the exhaustive search). For some test sequences such as 'Uli', 'Race1', 'Flamenco2', and 'Breakdancers', SA can converge to the global minima rapidly. For other sequences, the approximate results from the annealing process were very close to the global optima. As shown in Table 1, the difference was less than 0.2%.

Table 1 Simulation results for the simulated annealing (SA) algorithm

Test sequences	Number of views	Global minima	Best results	Worst results	Minimum number of iterations	Maximum number of iterations
Ballroom	8	743944	744430 (0.065%)	747970 (0.54%)	120	240
Exit	8	456090	456940 (0.19%)	458330 (0.49%)	120	216
Uli	8	2658700	2658700 (0.0%)	266900 (0.39%)	96	192
Race1	8	937980	937980 (0.0%)	946660 (0.93%)	144	216
Flamenco2	5	345830	345830 (0.0%)	347290 (0.42%)	75	120
Breakdancers	8	1248200	1248200 (0.0%)	1250120 (0.15%)	144	216
Rena	16	623520	624470 (0.15%)	628240 (0.76%)	336	480
Akko&Kayo	15	924230	926090 (0.2%)	931460 (0.78%)	315	450

The percentages in best and worst results columns denote the differences from the global minima

We used the best and worst results from SA to construct the inter-view prediction structures. The optimal prediction structures were quite different from the reference prediction structures of MVC. The best and worst results were also quite different from each other. The optimal references were not always the nearest neighbor views. Figs. 7 and 8 illustrate the prediction structures for sequences 'Ballroom' and 'Flamenco2', respectively. We compared the coding performance of the proposed solution with that of JMVM software. In temporal dimension, the proposed solution used the same GOP length and the same hierarchical B scheme as JMVM. In inter-view dimension, the optimal prediction structure obtained was used. GOPSize and IntraPeriod were set to 8 and three fixed QP settings specified in MVC common test conditions (Su *et al*., 2006) were used. For the sequence 'Uli', the correlation between views was so low that little coding gain can be expected from inter-view prediction. For other sequences, however, especially for those not arranged in 1D camera arrays, the proposed solution has exploited the inter-view correlation to the best of its potential. Fig. 9 shows the coding results of the test sequences. Consequently, the optimal prediction structure for MVC outperformed the reference prediction structure of JMVM by 0.1–0.8 dB PSNR gains. The coding performances of prediction structures based on the worst results were sometimes quite close to the optimal ones (e.g., for sequences 'Flamenco2', 'Breakdancers', 'Rena', and 'Akko&Kayo'). In general, 0–0.2 dB PSNR differences were observed between the best and worst results of SA. For 'Exit', the performance of 'OP_W' was close to JMVM. For other sequences, still 0.1–0.8 dB PSNR gains were obtained from 'OP_W'.

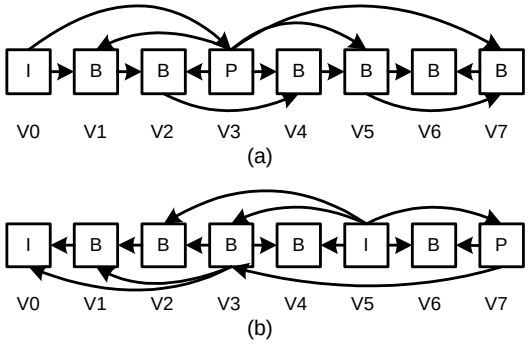

Fig. 7 Proposed prediction structure from our model for sequence 'Ballroom'

(a) Best result, coding order: {V0, V3, V1, V2, V4, V5, V7, V6}; (b) Worst result, coding order: {V5, V7, V3, V2, V1, V4, V6, V0}

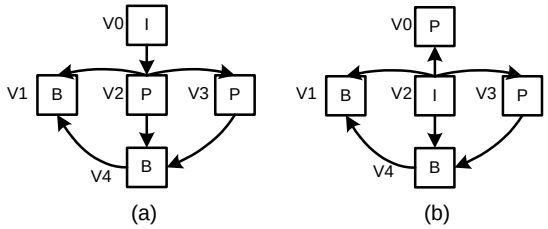

Fig. 8 Proposed prediction structure from our model for sequence 'Flamenco2'

(a) Best result, coding order: {V0, V2, V3, V4, V1}; (b) Worst result, coding order: {V2, V3, V0, V4, V1}

5 Discussion

Direct searching for the prediction structure is difficult to implement because the process starts from a complicated graph and also results in a complicated graph. Without B pictures, the searching process can be implemented using the minimum spanning tree method to generate an optimal tree from a directed complete graph. B pictures are efficient, however, in prediction structure and should be counted. In the proposed model, the design of prediction structure is converted to the arrangement of coding order so that B pictures can be dealt with. Nevertheless, it is an NP-hard problem (Cormen *et al*., 2001) to find the best coding order from the coding order domain. For a view number of N, it will take $N!$ iterations to perform an exhaustive search. When N=5, there are only 120 possible coding orders in the domain. As N increases, however, the amount of possible results as well as the computational complexity increases rapidly. When N=8, 40320 iterations are needed to perform an exhaustive search, and when N=16, the number of iterations increases to 2.09×10^{13}.

In our solution, the SA algorithm is employed as an effective method for finding these results. As SA is a stochastic algorithm, we also use the exhaustive search to generate the global optima as the reference. In comparison, SA greatly reduces the computational complexity whereas the results are close to the global optima. When N=5, SA converges in 75–120 iterations, which is comparable to the processing time of exhaustive search. When N=16, however, only less

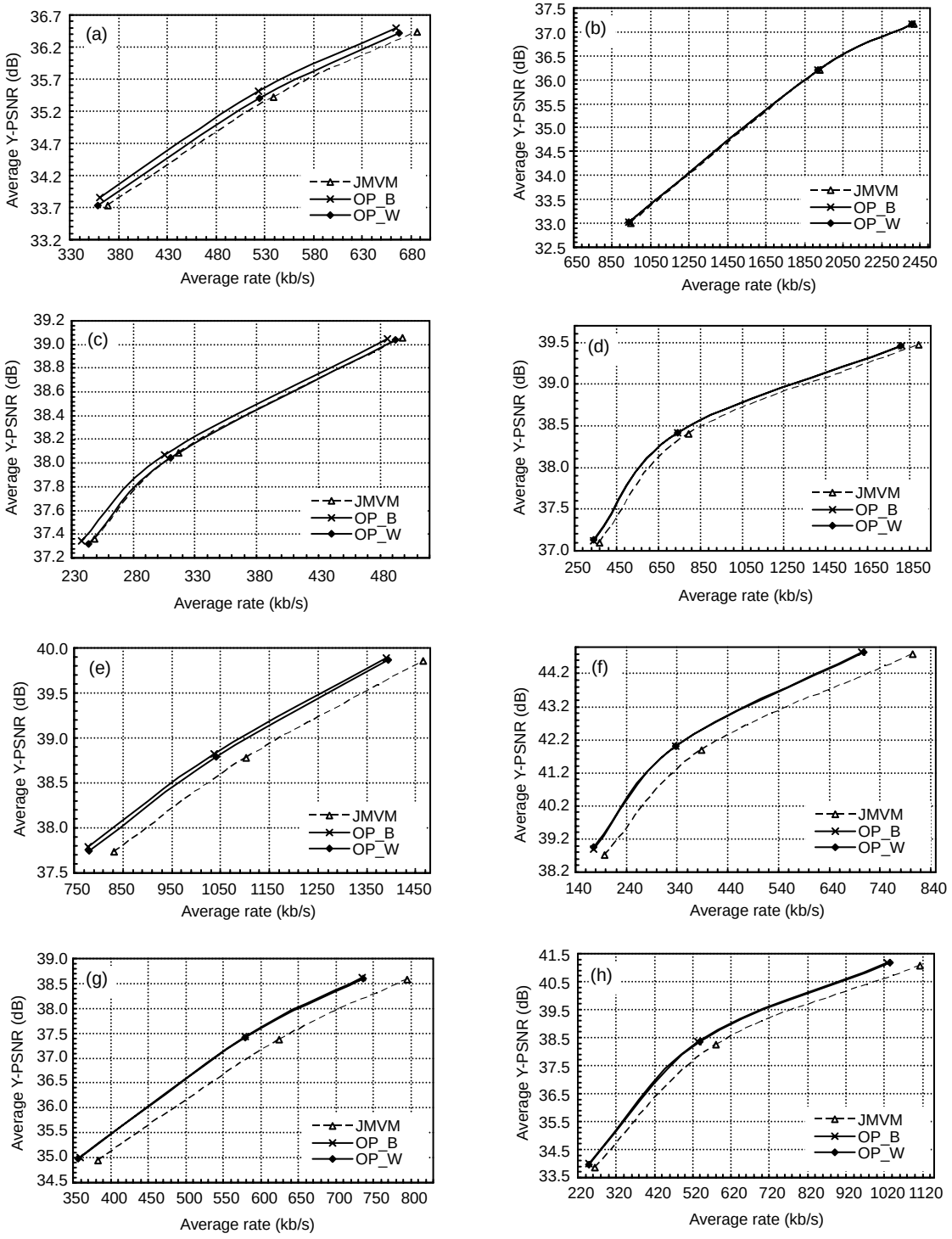

Fig. 9 Coding results for the proposed and reference prediction structures for multi-view video coding (MVC)
JMVM: JMVM reference prediction structure; OP_B: the optimal prediction structure based on the best result of SA; OP_W: the prediction structure based on the worst result. (a) Ballroom: 640×480 @ 25 frames/s; (b) Uli: 1024×768 @ 25 frames/s; (c) Exit: 640×480 @ 25 frames/s; (d) Breakdancers: 1024×768 @ 15 frames/s; (e) Race1: 640×480 @ 30 frames/s; (f) Rena: 640×480 @ 30 frames/s; (g) Flamenco2: 640×480 @ 30 frames/s; (h) Akko&kayo: 640×480 @ 30 frames/s

than 500 iterations are needed for SA. The processing time is reduced by 11 orders of magnitude. It is also observed that the number of SA iterations increases very slowly as N grows. The proposed method is obviously applicable to an arbitrary number of views.

This work is concentrated on the optimization of searching process. We assume that the coding costs are given and adopt the output bits at a predefined distortion as the measure of coding costs. Indeed, determining the cost table is an interesting subject. For N=8, each view needs one I cost, 7 P costs (reference from other 7 views), and 21 B costs (two references chosen from other 7 views). There will be 8×(1+7+21)=232 cost computations. A possible solution to reducing the complexity is to build a mathematical model related to the camera parameters. Furthermore, there is no scenario change in the MVC test sequences. Thus, we can use the coding costs from the first few time-points to decide the prediction structure for the whole sequence. In case of scenario changes, since camera parameters are independent of the scenario, the mathematical model related to camera parameters can be combined with the proposed SA method and used to avoid re-optimization. This is the subject of future investigations.

6 Conclusions

This paper proposes a novel method to optimize the inter-view prediction structures for MVC. The construction of prediction structure is simplified to the arrangement of coding order. Then, a simulated annealing algorithm is employed to minimize the total cost in order to obtain the best coding order. Experiment results show that the annealing process converges rapidly to satisfactory states. The SA algorithm exploits the inter-view correlation to the best of its potential, and the optimal prediction structures generated achieve 0.1–0.8 dB higher PSNR performance than the reference prediction structure of JMVM. The proposed method is applicable to arbitrary camera arrangements.

References

Bohachevsky, I.O., Johnson, M.E., Stein, M.L., 1986. Generalized simulated annealing for function optimization. *Technometrics*, **28**(3):209-217. [doi:10.2307/1269076]

Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C., 2001. Introduction to Algorithms (2nd Ed.). MIT Press, Cambridge, USA, p.966-1021.

Kalva, H., Christodoulou, L., Mayron, L., Marques, O., Furht, B., 2006. Challenges and Opportunities in Video Coding for 3DTV. IEEE Int. Conf. on Multimedia and Expo, p.1689-1692. [doi:10.1109/ICME.2006.262874]

Kang, J.W., Cho, S.H., Hur, N.H., Kim, C.S., Lee, S.U., 2007. Graph Theoretical Optimization of Prediction Structure in Multiview Video Coding. IEEE Int. Conf. on Image Processing, p.429-432. [doi:10.1109/ICIP.2007.4379613]

Kirkpatrick, S., Gelatt, C., Vechhi, M., 1983. Optimization by simulated annealing. *Science*, **220**(4598):671-680. [doi:10.1126/science.220.4598.671]

Laarhoven, P., Aarts, E., 1988. Simulated annealing: theory and applications. *Math. Its Appl.*, **12**(1):108-111.

Li, D.X., Zheng, W., Xie, X.H., Zhang, M., 2007. Optimising inter-view prediction structure for multiview video coding with minimum spanning tree. *Electron. Lett.*, **43**(23): 1269-1271. [doi:10.1049/el:20072465]

Merkle, P., Smolic, A., Muller, K., Wiegand, T., 2007. Efficient prediction structures for multiview video coding. *IEEE Trans. Circ. Syst. Video Technol.*, **17**(11):1461-1473. [doi:10.1109/TCSVT.2007.903665]

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, **21**(6):1087-1090. [doi:10.1063/1.1699114]

Muller, K., Merkle, P., Schwarz, H., Hinz, T., Smolic, A., Wiegand, T., Oelbaum, T., 2006. Multi-view Video Coding Based on H.264/MPEG4-AVC Using Hierarchical B Pictures. Picture Coding Symp., p.385-390.

Schwarz, H., Marpe, D., Wiegand, T., 2006. Analysis of Hierarchical B Pictures and MCTF. IEEE Int. Conf. on Multimedia and Expo, p.1929-1932. [doi:10.1109/ICME.2006.262934]

Smolic, A., Mueller, K., Merkle, P., Fehn, C., Kauff, P., Eisert, P., Wiegand, T., 2006. 3D Video and Free Viewpoint Video—Technologies, Applications and MPEG Standards. IEEE Int. Conf. on Multimedia and Expo, p.2161-2164. [doi:10.1109/ICME.2006.262683]

Su, Y., Vetro, A., Smolic, A., 2006. Common Test Conditions for Multiview Video Coding. ITU-T SG16/Q6. Doc. JVT-T207, Klagenfurt, Austria.

Zomaya, A.Y., 2001. Natural and simulated annealing. *Comput. Sci. Eng.*, **3**(6):97-99. [doi:10.1109/MCISE.2001.963434]