# Reduced Resolution Method: A Visual Signal Coding Technology for Mobile Communications

●Takashi Itoh

*(Manuscript received September 20, 2002)*

**The videophone service has been a dream for a long time, but there have been several reasons that prevented the service from becoming a reality. These are 1) the available data transmission bit rates were not high enough, 2) there were no encoding algorithms suitable for error prone transmission environments, and 3) the signal processing power of LSIs was too low. However, with the start of third-generation mobile communication services and the development of relevant technologies, such obstacles have almost disappeared and products with videophone capability are already on the market. This paper describes a visual signal encoding algorithm that Fujitsu has developed for the MPEG-4 system and describes its effectiveness under error free and error prone conditions. On average, our algorithm showed a 1 dB higher picture quality and an 18% smoother motion compared to the simple MPEG-4. The new algorithm is called the reduced resolution method.**

## 1. Introduction

To realize practical mobile audiovisual communication in third-generation mobile systems, an efficient video-encoding algorithm that can cope with the lower transmission bit rates and higher data error rates of mobile environments is essential (**Figure 1**). The ISO/IEC MPEG-4 standard is one of the solutions to these problems. Compared to previous algorithms, it produces a better picture quality at low bit rates (e.g., 64 kb/s) and takes tolerance to data errors into account.

The amount of information needed to transmit a video sequence changes dynamically from moment to moment. If part of a video sequence is stationary, almost no information needs to be sent. On the other hand, if the scene is constantly changing, a large amount of information needs to be sent. If the capacity of the transmission channel is only, for example, 64 kb/s, this wide variation in the amount of information to be sent results in an irregularity in the rate at which frames are sent.

To cope with this problem, conventional encoding algorithms, including MPEG-4, mainly control the quantization step size, which decreases the reproduced picture quality. We have therefore proposed a dynamic resolution control method for the prediction error signal to keep this defect as low as possible.

In the method, a reduced resolution encoding is carried out when the input scene is highly active. In the new method, the rate at which frames are encoded is averaged so that users can enjoy a higher picture quality. In addition, the method improves immunity to the high rate of data errors that inevitably occur in a mobile environment.

Section 2 of this paper briefly describes the MPEG standardization activities. Section 3 introduces our new method, including simulation results that show its effectiveness. Lastly, Section 4 describes the performance of the method when it is used with an error prone environment.

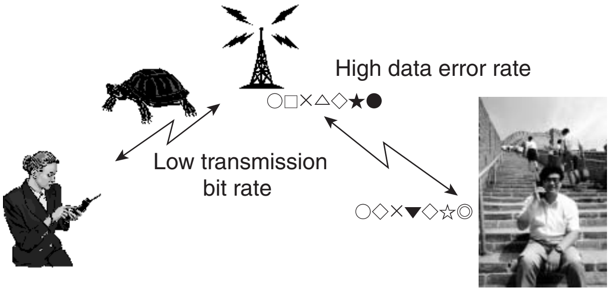



Figure 1
Mobile communication system with video.

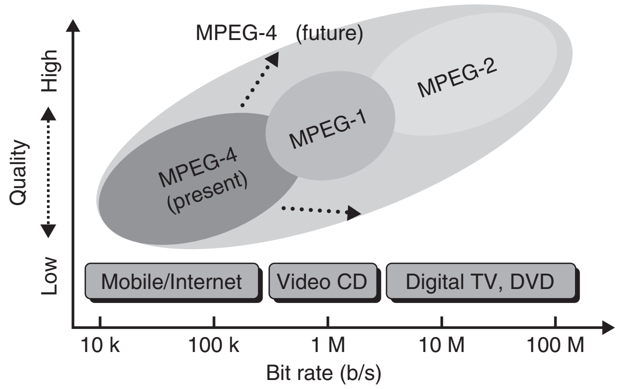



Figure 2
MPEG standards.

## 2. Standardization activities

One of the obvious requirements for using encoded data is a standardized encoding algorithm.[note)] Two well-known standardization organizations are the ITU-T and ISO/IEC. MPEG is the name of a group under ISO/IEC's JTC1 (Joint Technical Committee 1) and is also the name of the standards that the MPEG group has made.

### 2.1 MPEG series

The relationship between the MPEG-1, MPEG-2, and MPEG-4 standards is shown in **Figure 2**.

MPEG-1 was standardized in 1992. Its main objective was to store a one-hour movie on a CD-ROM. However, CD-ROMs not only have a low capacity, but their data reading speed is set to a basic 150 kbytes/s.

MPEG-2 was standardized in 1994 and includes HDTV in its cover range. MPEG-2 is a well-known standard and is used in digital satellite broadcasting and DVDs.

Both MPEG-1 and MPEG-2 have been developed for use with rather higher bit rates such as 1.2 Mb/s and 6 Mb/s. Furthermore, they do not take errors into account, so errors in storage media and/or transmission errors need to be solved by some means outside of the MPEG standards.

note) The target of recent standards is not algorithms themselves, but the definitions of expressions of encoded output data.

### 2.2 MPEG-4

MPEG-4 was standardized in 1998. At the beginning of the standardization activity, the target was to develop an encoding algorithm that had a much higher efficiency compared to that of conventional algorithms. After several meetings of experts, the target of the standardization was changed to develop new features which conventional algorithms did not have, for example, applications under mobile environments. In MPEG-4, more than 10 profiles are defined, each having different functions for applications in a wide variety of services.

## 3. Reduced resolution method

Recent efficient video encoding algorithms reduce the amount of information in the input visual signals by analyzing the correlation between frames. This method is called interframe prediction. When this method is used, only the differences from the one or more reference frames are transmitted. This means that the amount of information generated per frame is large when the correlation is small, for example, when there are many changes in a video sequence or the camera moves to another scene. When this occurs, the amount of information can easily become three or four times the average value. This makes motion in the reproduced video less smooth and therefore of a lower quality. We have therefore developed a reduced resolution method to solve this problem.

## 3.1 Conventional algorithm

The interframe prediction method is shown in **Figure 3**. The current frame being encoded is compared with the previously encoded frame to extract the difference between them. Then, to get the current frame at the receiver, the difference is transmitted and added to the previous frame, which is stored in the receiver. The minus sign near the car at the lower-left position means that the car needs to be eliminated at the receiver side to get the current frame.

This method cannot cope when a person or object moves within a frame, and a motion estimation method is widely used to compensate for such motions.

## 3.2 The reduced resolution method

The essence of our reduced resolution method is shown in **Figure 4**.

In a video sequence depicting a highly active scene, first a prediction error is detected by calculating the difference between the frame being encoded and the reference frame. Then:

1) The resolution of the prediction error is down-sampled to one half, both horizontally and vertically. The one-half sampling ratio was selected for the following reasons:
   - QCIF (Quarter Common Intermediate Format), which is a typical resolution defined in MPEG-4, is one-quarter of CIF (Common Intermediate Format), which is another typical MPEG-4 resolution.
   - A digital filter can be realized rather simply at a one-half sampling ratio.
   - Calculation for adjusting motion vectors is easy at a one-half sampling ratio.
2) A reduced-size prediction error is transmitted.
3) The received reduced-size prediction error is up-sampled at the receiver side to get the original-size error data.
4) The up-sampled prediction error of the original size is added to the predicted picture (not shown in the figure) to produce the reproduced picture.

The merits of this method are:

1) The amount of information generated in a highly active scene can be reduced. Hence, the frame rate can be kept high.
2) The resolution of the unchanged portion in a picture is kept high. Since the sensitivity of human eyes is high in still areas, the subjective picture quality becomes better.

The demerits are:

1) A calculation for deciding the resolution must be added at the encoder side. The type of

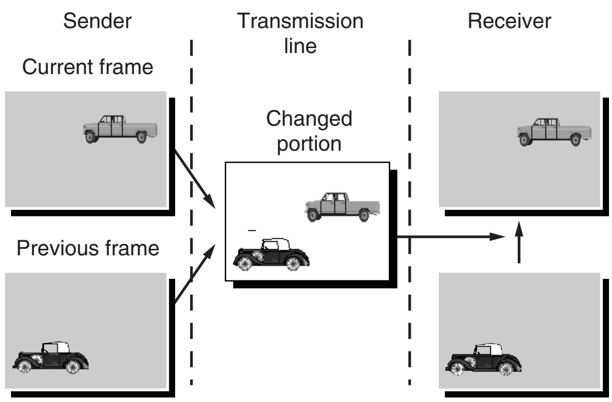



Figure 3
Conventional encoding algorithm.
- Interframe prediction -

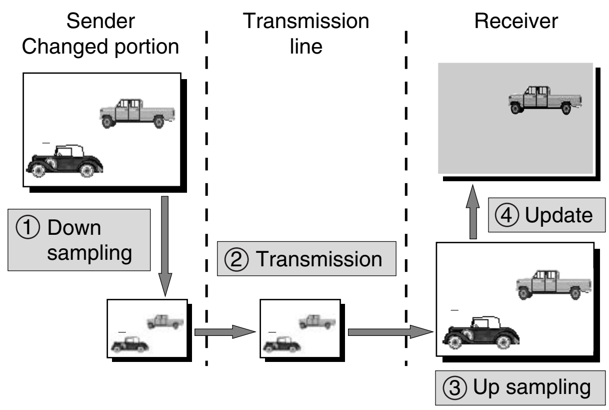



Figure 4
New encoding algorithm.
- Reduced resolution method -

calculation depends on the motion estimation algorithm that is used.

2) To ensure an improvement in coding performance, the decision algorithm must be carefully optimized.

The details of the reduced resolution method are described in Reference 1).

### 3.3 Simulation for performance evaluation

We have conducted some computer simulations to test the effectiveness of the method. Some of our results are shown in **Figure 5**.

The sequence used is the Bowing sequence, which is one of the test sequences used in MPEG-4 standardization activities. In the sequence, a person comes into the screen, bows, and goes out of the screen. The bit rate was set to 48 kb/s because the total bit rate of a typical mobile communication service is 64 kb/s and bits for audio and control signals need to be subtracted from this.

Figure 5 (a) shows the number of generated bits per frame for an input of 30 frames per second. The thick line stands for the reduced resolution method, and the thin line stands for the conventional CIF size encoding method. Because the selection of encoding frames depends on the amount of information generated for the previously encoded frames, it is not guaranteed that the same frame is encoded in both of the two methods. As is clearly seen, there are four peaks in the conventional method's curve, which make the motion less smooth because fewer frames can be encoded per second. Our method, on the other hand, succeeds in keeping the number of generated bits almost constant. This results in smoothly reproduced motion pictures.

The lower part of Figure 5 (a) shows the resolution. As can be seen, the lower resolution was selected automatically at the four peaks.

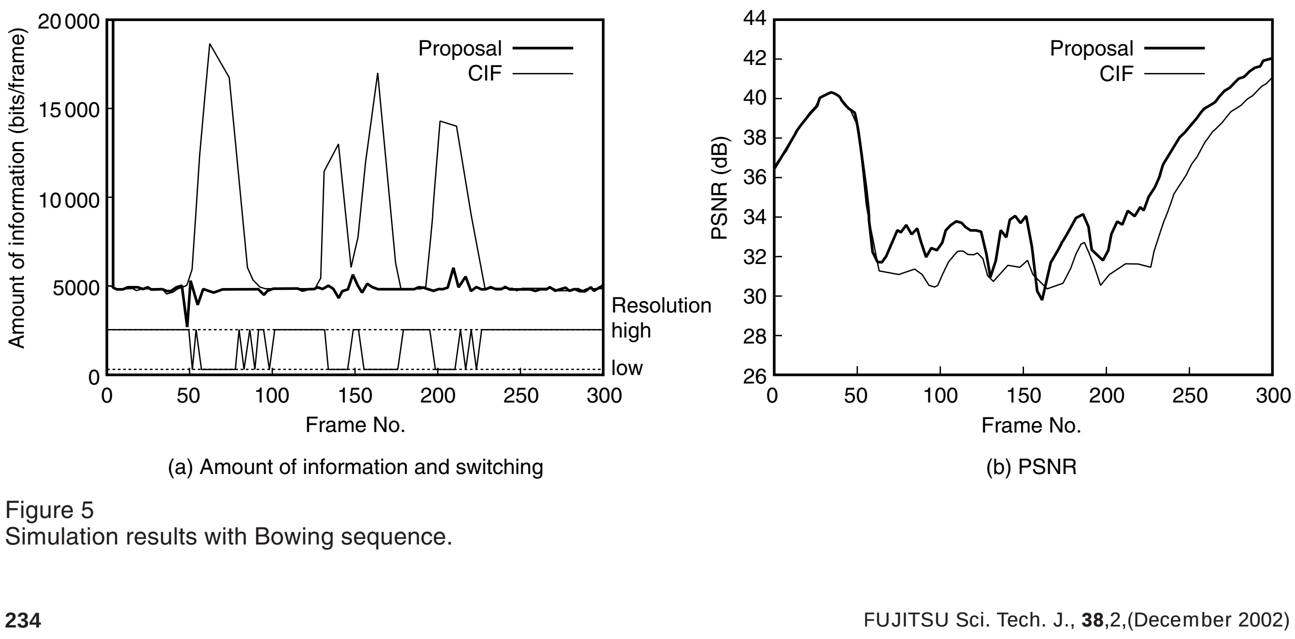
Figure 5 (b) shows the PSNR (Peak Signal to Noise Ratio) of the luminance component. The vertical axis is the picture quality, so a higher value on this axis indicates a higher quality. Again, our method shows better performance with fewer bits assigned.

Subjectively, our method provides obviously better quality pictures, and the reason for this is that it transmits at an almost constant frame rate.

## 4. Error resilience

The reduced resolution method is expected to have a preferable effect when used with an error prone transmission environment. It transmits more frames with less information on average compared with CIF size encoding, and therefore:

1) The number of erroneous bits in a picture




(a) Amount of information and switching

(b) PSNR

Figure 5
Simulation results with Bowing sequence.

decreases, so the quality of each picture becomes better.

2) The number of pictures transmitted per unit of time (e.g., every second) increases, and therefore error recovery can be achieved faster than with a low frame rate transmission.

In mobile communication services, the transmission data error rate is much higher than in conventional wired services: a typical bit error rate (BER) in a mobile communication service can be up to $10^{-3}$. Detailed data about the performance of our method under such severe conditions is given in Reference 2).

## 4.1 Intra-refresh

Because the reduced resolution method itself has no error recovery capability, we adopted an intra-refresh method that encodes part of a picture without referring to other pixels. If many parts of a picture are encoded with the intra-refresh mode, the recovery time from an error can be shortened. However, this approach decreases the number of transmitted pictures per unit of time since the correlation between frames is not utilized and therefore the amount of information in each picture increases. Therefore, it becomes important to find the optimal size of intra-coding area by taking both the error recovery time and the transmitted picture rate into account.

## 4.2 Error-free condition

We carried out computer simulations to obtain some insights about the optimal size of the intra-coding area.

All macroblocks (i.e., $16 \times 16$ luminance pixels) must be encoded with the intra-coding mode at least every 132 pictures to prevent the computational error from accumulating. By adjusting this number, the size of intra-coding area in each picture can be controlled.

Under error free conditions, the intra-coding mode essentially need not be used. In addition, the intra-coding mode consumes more bits to express the same picture than other coding modes such as interframe prediction and motion compensated prediction. Therefore, it is predicted that the maximum intra-coding interval of 132 (i.e., the least frequent refresh) will show the best performance.

To examine the effect of the reduced resolution method, average PSNR values are calculated for several intra-refresh intervals. **Figure 6** shows the intra-refresh interval versus the average PSNR. The solid line with white circles shows the performance of the reduced resolution method, while the dotted line with black circles shows the result of the CIF size encoding method. Our method has about a 1 dB higher performance than the CIF size method over the entire range of refresh intervals. In both methods, a less frequent refresh means a better PSNR, which is what we expected before the simulation.

Another aspect of picture quality enhancement is the smoothness of the reproduced video. **Figure 7** shows the number of encoded frames. Here again, the reduced resolution method has about an 18% better performance compared to the CIF size method.

With the higher quality of each picture shown in Figure 6 and the higher frame rate shown in Figure 7, the subjective picture quality of our method is better than that of the CIF size method.

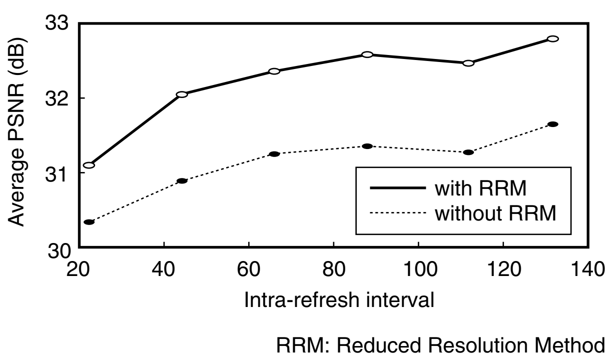



Figure 6
Average PSNR of luminance signal vs. intra-refresh interval (error free).

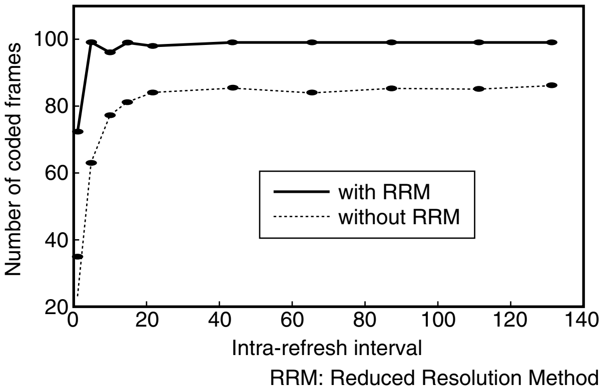



Figure 7
Number of encoded frames vs.
intra-refresh interval (error free).

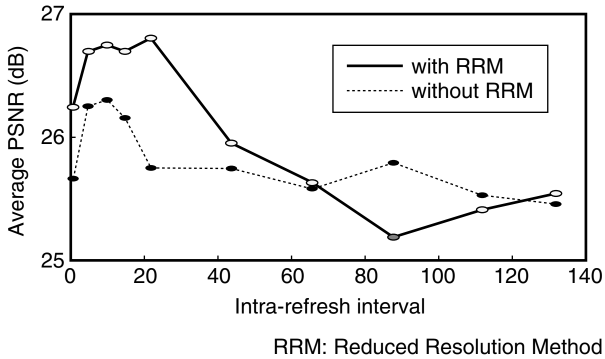



Figure 8
Average PSNR of luminance signal vs.
intra-refresh interval (BER = $10^{-3}$ ).

## 4.3 Error prone condition

Under error prone conditions, however, longer intra-coding intervals may cause a longer recovery time after a data error since data that is independent from other pixels is transmitted the least frequently. Too short an interval, on the other hand, makes the size of the intra-coded area in a picture too large, and each picture has more information. This results in an increase in the number of erroneous bits in a picture and a decrease in the number of encoded pictures per second. These two changes degrade the quality of the reproduced pictures. Therefore, the best interval must be investigated.

The simulation procedure is as follows:

- Encode input pictures according to MPEG-4 while varying the intra-refresh interval.
- Inject an error pattern with a bit error rate of $10^{-3}$ into the encoded data.
- Decode the erroneous data, and calculate the PSNR.

Here, the injected error pattern was used with a shift operation to attack a different part of the encoded data. As a result, many PSNRs were obtained for each refresh interval. The average of these PSNRs is used in the following discussion.

From **Figure 8**, which corresponds to Figure 6 in the error free case, longer refresh intervals such as 66 and 132 provide lower performances. On the other hand, frequent intra-refresh rates such as 22 or less tend to have higher PSNR values. However, the effect of the reduced resolution method is not so simple. In the frequent refresh area, the advantage of the method is clear, but since the influence of transmission errors becomes dominant in the less frequent refresh area, both coding methods show poor PSNR values.

The number of encoded pictures is the same as in Figure 7 since errors are injected after encoding.

Taking Figures 7 and 8 into account, an intra-refresh interval of 22 seems the best value.

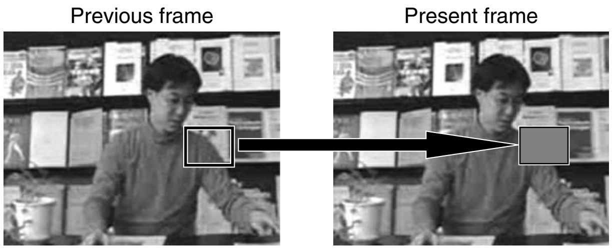



Figure 9
Error resilience of MPEG-4: Error concealment.

## 4.4 Alternative error concealment method

In some cases with errors, it is impossible to reproduce a transmitted picture entirely or even in part. In **Figure 9**, the rectangle on the right could not be decoded. The simplest way to recover the lost rectangle is to copy the same size of rectangle at the same position from the previously decoded picture. However, a more sophisticated method is applicable which searches for a better rectangle from the previous frame based on the similarity of the surrounding correctly decoded blocks and/or the motion vectors of the surrounding blocks.

## 5. Conclusion

Among the third-generation mobile communication services, the videophone is one of the most anticipated ones. To realize this service, MPEG-4 has been adopted as the visual signal encoding algorithm. The simple MPEG-4 has a good performance at lower bit rates such as 64 kb/s. However, compared to the simple MPEG-4, our reduced resolution method can provide better service quality, both in terms of the picture quality of each frame and in the smoothness of the reproduced pictures' motion. Specifically, compared to the simple MPEG-4, our method provides about a 1 dB better picture quality and about an 18% better smoothness. The picture quality provided by our method can be considerably enhanced by combining it with the intra-refresh coding and error concealment technologies we have developed. We will evaluate the quality of the pictures obtained through this combination in further detail in a future study.

One of the urgent priorities is to seek killer applications that utilize motion pictures as a core media for their success. In addition, further evaluation and decision-algorithm enhancement may be necessary to attain a better performance with various motion sequences.

## References


1) A. Nakagawa et al.: Dynamic Resolution Conversion Method for Low Bit Rate Video Transmission. *IEICE Trans. Commun.*, **E84-B**, 4, p.930-940 (April 2001).
2) T. Itoh et al.: Evaluation of error robustness of the reduced resolution method for low bit rate video transmission. ICCC99.


**Takashi Itoh** received the B.E. degree in Chemical Engineering and the M.E. degree in Information Engineering from the University of Tokyo, Tokyo, Japan in 1981 and 1983, respectively. He joined Fujitsu Laboratories Ltd., Kawasaki, Japan in 1983, where he has been engaged in research and development of algorithms, LSIs, and CODECs for visual signal coding and research and development of visual communication systems. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan, the Information Processing Society of Japan, and the Institute of Electrical Engineers of Japan.