(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 April 2005 (21.04.2005)

PCT

(10) International Publication Number
WO 2005/036893 A2

(51) International Patent Classification[7]: H04Q

(21) International Application Number:
PCT/KR2004/002569

(22) International Filing Date: 8 October 2004 (08.10.2004)

(25) Filing Language: Korean

(26) Publication Language: English

(30) Priority Data:
10-2003-0070499 10 October 2003 (10.10.2003) KR

(71) Applicants *(for all designated States except US)*: **ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE** [KR/KR]; 161, Gajeong-dong, Yuseong-gu, Daejon-si 305-350 (KR). **INTERJUNGBO CO., LTD** [KR/KR]; 999-6, Daechi-dong, Gangnam-gu, Seoul 135-280 (KR). **RESEARCH AND INDUSTRIAL COOPERATION GROUP** [KR/KR]; 119 Munjiro Yuseong-gu, Daejon-si 305-714 (KR).

(72) Inventors; and

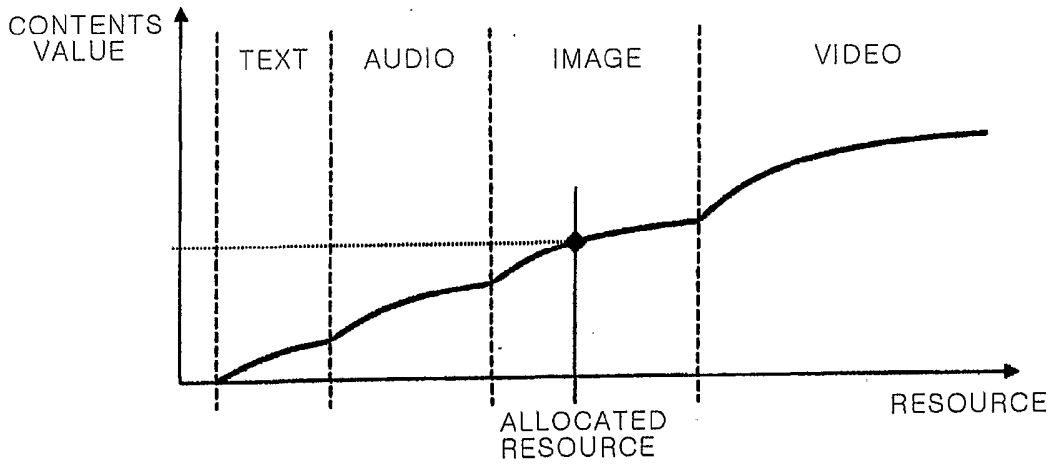
(75) Inventors/Applicants *(for US only)*: **RO, Yong-Man** [KR/KR]; 816-1102, Saemirae 8 Apt., Nouen-dong, Yuseong-gu, Daejon-si 305-325 (KR). **TRUONG, Cong, Thang** [KR/KR]; Image and video system Lab., Information and, communications University Educational Foundation, 103-6, Munji-dong, Yuseong-gu, Daejon-si, 305-714 (KR). **HONG, Jin-Woo** [KR/KR]; 130-702 Hanbit Apt., Euen-dong, Yuseong-gu, Daejon-si 305-755 (KR). **KIM, Jae-Gon** [KR/KR]; 203-402 Saemmeori Apt., Dunsan-dong, Seo-gu, Daejon 302-777 (KR). **NAM, Je-Ho** [KR/KR]; 119-33, Yonheel-dong, Seodaemun-gu, Seoul 157-909 (KR).

(74) Agent: **KIM, Myung-Shin**; Myung-Shin & Partners International, Patent & Trademark Office,, 12 Fl., Jindo Bldg., 37, Dowha-dong, Mapo-gu,, Seoul 121-040 (KR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR CONVERTING THE MODALITY OF MULTIMEDIA CONTENTS TO SUPPORT THE QUALITY OF SERVICE ACCORDING TO MEDIA RESOURCE



(57) Abstract: The present invention relates to modality conversion for supporting Quality of Service (QoS) according to media resources, which includes the steps of receiving a modality conversion descriptor in which the characteristics of the modality conversion of the multimedia contents are described, receiving the multimedia contents, and converting the modality of the multimedia contents into a modality determined according to a media resource and the modality conversion descriptor. According to the present invention, a systematic approach to a design for an overlap content value model is provided so that the conversion boundaries between modalities can be quantitatively calculated. As a result, the correlation between various modalities can be established in a single model, and modality conversion supporting the optimal QoS can be performed.

WO 2005/036893 A2

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHOD AND APPARATUS FOR CONVERTING THE MODALITY OF MULTIMEDIA CONTENTS TO SUPPORT THE QUALITY OF SERVICE ACCORDING TO MEDIA RESOURCE

Technical Field

The present invention relates to a method and apparatus for converting the modality of multimedia contents to support the quality of service according to media resources and, more particularly, to the modality conversion of multimedia contents to support the quality of service of the multimedia contents consumed in a ubiquitous computing environment.

Background Art

The case where a terminal does not support content resources of the same kind but accesses and consumes universal multimedia is a new trend in multimedia communication. In a system for Universal Multimedia Access (UMA), the adaptive conversion of contents is the most important process for supporting the Quality of Service (QoS) for a user. The adaptive conversion of contents may be considered from two points of views. One is content transcoding that changes the bit rate (or quality) of contents without converting the modality of the contents,

and the other is modality conversion that converts contents from one modality (e.g., video) to another modality (e.g., still image).

A content value model (or content value curve), which represents a relative correlation between a content value (or content quality) and a resource, has been discussed in some recent research. In a thesis "JPEG Compression Metric as a Quality Aware Image Transcoding" published in 2nd Symp. Internet Technologies and Systems" by S. Chandra and C. S. Ellis in October of 1999, and a thesis "JPEG Compression Metric as a Quality Aware Image Transcoding, 2nd Symp. Internet Technologies and Systems," Boulder, CO: USENIX, Oct. 1999" by S. Chandra and C. S. Ellis, wide research into quality measurement specific to JPEG images was performed based on a compression rate closely related to the amount of resource. In a thesis "Adapting Multimedia Internet Content for Universal Access, IEEE Trans. Multimedia, Vol. 1, No. 1, pp. 104-114, Mar. 1999" published in IEEE Trans. Multimedia, Vol. 1, No. 1, pp. 104-114 by Rakesh Mohan, John R. Smith, and Chung-Sheng Li, the case where a correlation between a content value and a resource is represented using a single concave function is discussed when a content representation design includes content versions having different modalities. However, the concave function does not properly represent a correlation between content values having different modalities.

Currently, most content values for QoS must be calculated according to specific modalities, respectively. The reason for this is that each modality has a specific characteristic and the qualities of different modalities may be measured in different dimensions. However, a systematic method of easily performing modality conversion and content scaling needs the combination of the content values for different modalities.

Intuitively, if the characteristics of contents for several resources, that is, restrictions, are given, a service provider (down) scales the contents to be suitable for the characteristics for the resources until the best QoS is provided to the user. However, in some cases, the quality of the scaled contents may not be suitable for the user. In this case, a solution to the above problem is to convert the modality of the contents into another modality. For example, when a bandwidth is excessively narrow, the transmission of a series of different images may be more suitable than the streaming and transmission of low-quality video. Such a method is a typical example of the conversion of a modality from video into images.

When viewed from a QoS point of view, the most important issue in the modality conversion is "which resource characteristic requires that a current modality is converted into a different modality." Most current systems supporting modality conversion operate only when terminals

do not support specific modalities. However, when terminals support modalities but resources are restricted, any prior art inventions have not provided a systematic solution to modality conversion guaranteeing the best QoS. The present invention provides a method of finding modality conversion boundaries for the above solution based on QoS for modalities.

Disclosure of the Invention

Accordingly, an object of the present invention is to provide a time point and method of modality conversion, which guarantees the best QoS from a QoS point of view in different modalities as well as in the same modality under various content and device environments.

In relation to such an object, the present invention proposes (1) a method of designing an overlap content model for QoS using the scale factors for different content modalities, and (2) a description of the QoS of modality conversion into eXtended Markup Language (XML) for the automation of modality conversion.

The present invention provides a method of converting a modality of multimedia contents to support QoS of the multimedia contents according to media resources, including the steps of receiving a modality conversion descriptor in which characteristics of modality conversion of the

multimedia contents are described, receiving the multimedia contents, and converting the modality of the multimedia contents into a modality that is determined according to a media resource and the modality conversion descriptor.

The modality conversion descriptor describes the content value curves and scale factors for the modalities of the multimedia contents. Each of the content value curves is obtained by combining content value curves measured according to two or more different qualities.

Preferably, the modality conversion step includes the steps of obtaining conversion boundaries using the content value curves and the scale factors for the modalities, determining the optimal modality for the media resource using the conversion boundaries, and converting the multimedia contents into the determined optimal modality.

Furthermore, the present invention provides an apparatus for converting a modality of multimedia contents to support QoS of the multimedia contents according to media resources, including a means for receiving a modality conversion descriptor in which characteristics of modality conversion of the multimedia contents are described, and a means for converting the modality of the multimedia contents into a modality that is determined according to a media resource and the modality conversion descriptor.

An approach proposed by the present invention is composed of an overlap content value model that represents

the correlation between content values and resources to support QoS, and a design for the overlap content value model. In the present invention, contents are transmitted under heterogeneous network environments and consumed by various terminals. First, a data description structure, in which the characteristics of modality conversion of the multimedia contents to guarantee QoS are described, is received. Thereafter, the modality of the multimedia contents is converted into a 'modality for providing QoS according to the input data description structure.

Preferably, when the modality of the multimedia contents is converted to provide QoS according to the input data description structure, the modality of the multimedia contents is converted using the conversion boundaries between different content modalities. The step of converting the modality of the contents using the conversion boundaries between the different content modalities includes the steps of (A) finding the content value curves (modality curves) for modalities that represent correlations between content values and resources, (B) finding the scale factors for the modalities that represent the importance of the modalities, (C) mapping the modality curves to a single overlap content value model according to the scale factors, (D) obtaining conversion boundaries that are the intersection points of the modality curves in the overlap content value model, and

(E) converting a current modality into a different modality using the obtained modality conversion boundaries to support QoS under a restriction.

At the step of finding the content value curves (modality curves) of the modalities that represent the correlations between the content values and the resources, the content value can be measured according to different qualities (PSNR, MOS, etc.). When the content value is related to various resources, the modality curves are changed to modality surfaces.

At the step of finding the scale factors of the modalities that represent the importance of the modalities, the scale factors are automatically or manually provided by a content author or provider. At the step of mapping the modality curves to the single overlap content value model according to the scale factors, the content value is obtained from a combination of values measured according to different qualities.

Brief Description of the Drawings

FIG. 1 is a block diagram showing an apparatus for converting the modality of multimedia contents in accordance with an embodiment of the present invention;

FIG. 2 is a view showing the concept of the overlap content value model for content items in accordance with

the present invention;

FIG. 3 is a conceptual view showing the final content value function for content items in accordance with the present invention;

FIG. 4 is a conceptual diagram showing the construction (mapping) from two quality curves into a single final modality curve for a video content modality, in which FIG. 4a is a graph showing a quality curve measured at a Peak Signal to Noise Ratio (PSNR), FIG. 4b is a graph showing a quality curve measured at a Mean Opinion Score (MOS), and FIG. 4c is a graph showing a combined single final modality curve in the case where each scale factor is one; and

FIG. 5 is a flowchart illustrating a process of converting the modality of contents to support QoS.

Best Mode for Carrying Out the Invention

Hereinafter, with reference to the attached drawings, embodiments of the present invention are described in detail. In the drawings, the same reference numerals are used to designate the same or similar components or signals.

FIG. 1 is a block diagram showing an apparatus for converting the modality of multimedia contents in accordance with an embodiment of the present invention. A

multimedia content and modality conversion descriptor receiving unit 102 receives multimedia contents and the modality conversion descriptor of the multimedia contents through a network, and provides the received multimedia contents and descriptor to a modality conversion descriptor extracting unit 104 and a multimedia content modality converting unit 108. The modality conversion descriptor extracting unit 104 extracts the modality conversion descriptor of the input multimedia contents and provides the modality conversion descriptor to the multimedia content modality converting unit 108. A media resource information providing unit 106 provides a restriction, which is related to the provision of the multimedia contents and is provided by the network or terminal, to the multimedia content modality converting unit 108. The multimedia content modality converting unit 108 determines the optimal modality for a media resource provided by the modality conversion descriptor and the media resource information, and converts the modality of the multimedia contents into the optimal modality. The multimedia contents, which are converted into a new modality, are provided to the network or terminal through an output unit 110.

An overlap content value model provided by the present invention can represent the correlations between the content values, resources and modalities. Contents can

be represented along with the overlap content value model that represents content values according to different content modalities.

FIG. 2 shows an example of a content value model for contents having a video modality. In FIG. 2, the content value curves for various modalities can be allocated by a content provider or an author, or through calculation in the terminal.

FIG. 3 is a view showing the concept of the final content value function for content items in accordance with the present invention. The intersection points of modality curves represent the conversion boundaries between modalities as indicated by the dotted lines of FIG. 3. On the basis of the conversion boundaries, modality conversion can be quantitatively determined to maintain the maximum allowed QoS. The present invention includes a method of representing various content modalities by the overlap model as shown in FIG. 3 and determining a modality conversion point based on QoS.

A single modality curve is formed by a content scaling operation adopted by a content scaler. For example, video contents are scaled to match the characteristics of a specific resource in consideration of space resolution, time resolution and coded bits per pixel. To form an overlap content model from such a single content model, a single curve including various modalities must be produced

on the single overlap content model.

It is assumed that the content value curve for the modality j of contents i is VM_{ij}. In this case, j = 1,· · · , J_i, and J_i is the number of modalities adoptable by the contents i. Furthermore, $VM_{ij} \geq 0$ must be fulfilled for all j=1,· · · , J_i, and j=1 is an index representing the original modality of the contents. In this case, the content value function for the contents i is expressed by the following Equation 1,

$$V_i = \max\{W_{ij} VM_{ij} | j = 1,...,J_i\} \quad (1)$$

In Equation 1, W_{ij} is a scale factor for the modality j of the contents i. In order to combine different modalities into a single model as proposed by Equation 1, the content author or provider automatically or manually allocates a proper scale factor W_{ij} to each of the modalities, so that the content values for different formalities must reflect the relative importance thereof, and have common measurement units for common content values.

Furthermore, the content value curve of a modality may be measured according to various qualities. For example, a video modality can obtain content value curves by calculating a PSNR and a MOS. Accordingly, to represent a content value curve measured according to a specific

quality, a term "quality curve" is used. As a result, a single modality curve may be formed of various quality curves.

$$VM_{ij} = \sum_{k} z_k \cdot VM^k{}_{ij} \qquad (2)$$

In Equation 2, $VM^k{}_{ij}$ is a quality curve measured according to the quality type k of contents. Furthermore, z_k is the scale factor of the quality type k to guarantee QoS. FIG. 4 is a graph showing an example of obtaining a modality curve, which shows a process of forming a single modality curve from two quality curves having differently measured content values (PSNR and MOS), respectively. The scale factor z_k is a parameter for combining the differently measured content values into a single modality curve, which can be previously provided by a content provider according to a QoS schedule, or be automatically calculated.

Meanwhile, when contents are related to various resources, modality curves are converted into modality surfaces in a multi-dimensional space. Furthermore, the above-described method proposed by the present invention can be applied and valuable regardless in spite of consideration to many resources. Collectively, in the modality conversion for QoS, the analytic characteristics of the overlap content value model can be summarized as

follows:

(1) The model includes various modality curves.

(2) The final content value function is composed of the upper parts of the overlap model.

(3) Each modality curve can be constructed by one or more content scaling operations.

(4) Each modality curve is characterized in that it is not reduced, and is not increased when the number of resources increases.

A summarized entire modality conversion process is shown in the flowchart of FIG. 5. Quality curves are calculated for n modalities using various measurement methods at step 10 and 11. The calculated quality curves for each of the modalities are combined into a single modality curve based on scale factors, which depend on quality types, using Equation 2 at step 20 and 21. In this case, as a result, n modality curves are obtained at step 30 and 31. After the scale factors for the modalities are roughly calculated, the modality curves are mapped to a single model based on the obtained scale factors so as to form a single overlap content value model at step 40. In the single overlap content value model for the contents (step 50), the intersection points of the modality curves can be found at step 60. A solution to under which

restriction a current modality should be converted into a different modality with QoS being supported can be obtained using the finally obtained modality conversion boundaries (step 70).

In forming the overlap content model in this system, metadata information representing parameters used needs to be represented in a structured form such as eXtended Markup Language (XML). Table 1 represents the modality conversion QoS technology, which is proposed by the present invention, in an XML form.

Table 1

```
<!-- ################################################ -->
<!-- Definition of ModalityConversionQoSType          -->
<!-- ################################################ -->

<complexType name="ModalityConversionQoSType">
    <complexContent>
        <extension base="dia:DIABaseType">
            <sequence>
                <element name="ModalityCurve" type="ModalityCurveType"
                    maxOccurs="unbounded"/>
            </sequence>
        </extension>
    </complexContent>
</complexType>

<complexType name="ModalityCurveType">
    <complexContent>
        <extension base="dia:DIABaseType">
            <sequence>
                <element name="Modality" type="mpeg7:controlledTerm" minOccurs="0"/>
                <element name="ModScale" type="float"/>
                <element name="QualityCurve" minOccurs="0" maxOccurs="unbounded">
                    <complexType>
                        <sequence>
                            <element name="UtilityRef" type="IDREF"/>
                            <element name="UtiScale" type="float"/>
                        </sequence>
                    </complexType>
                </element>
            </sequence>
        </extension>
    </complexContent>
</complexType>
```

In Table 1, the semantics of "ModalityCurveType" are defined as shown in Table 2,

Table 2

TITLE	DEFINITION
Modality	describe modalities of resources
ModScale	describe scale factors for modalities of resources
QualityCurve	describe content value curves for modalities of resources
UtilityRef	describe quality curves for modalities of resources
UtiScale	describe scale factors for quality curves

The above-described embodiments are used only for those skilled in the art to easily understand the present invention, but are not intended to limit the scope of the present invention. Those skilled in the art must note that various modifications and changes of the above-described embodiments are possible. In principle, the scope of the present invention is defined by claims described later.

Industrial Applicability

According to the construction of the present invention, a systematic approach to a design for an overlap content value model is provided so that the conversion boundaries between modalities can be quantitatively calculated. As a result, the correlation between various modalities can be established in a single model, and modality conversion supporting an optimal QoS can be

performed. Furthermore, the present invention is used as a basis for precisely determining modality conversion and content scaling, so that UMA is effectively achieved.

Claims

1. A method of converting a modality of multimedia contents to support Quality of Service (QoS) of the multimedia contents according to media resources, comprising the steps of:

receiving a modality conversion descriptor in which characteristics of modality conversion of the multimedia contents are described;

receiving the multimedia contents; and

converting the modality of the multimedia contents into a modality that is determined according to a media resource and the modality conversion descriptor.

2. The method according to claim 1, wherein the media resource is a network or terminal to which the multimedia contents are provided.

3. The method according to claim 1, wherein the modality conversion descriptor describes content value curves and scale factors for modalities of the multimedia contents.

4. The method according to claim 3, wherein the modality conversion step comprises the steps of:

obtaining conversion boundaries using the content

value curves and scale factors for the modalities;

determining an optimal modality for the media resource using the conversion boundaries; and

converting the multimedia contents into the determined optimal modality.

5. The method according to claim 4, wherein the conversion boundaries are values of the media resource corresponding to intersection points where the content value curves intersect with each other when the content value curves for the modalities overlap with each other according to the scale factors.

6. The method according to claim 3, wherein each of the content value curves is obtained by combining content value curves that are measured according to two or more different qualities.

7. An apparatus for converting a modality of multimedia contents to support QoS of the multimedia contents according to media resources, comprising:

means for receiving a modality conversion descriptor in which characteristics of modality conversion of the multimedia contents are described; and

means for converting the modality of the multimedia contents into a modality that is determined according to a

media resource and the modality conversion descriptor.

8. The apparatus according to claim 7, wherein the modality conversion descriptor describes content value curves and scale factors for modalities of the multimedia contents.

9. The apparatus according to claim 8, wherein the modality conversion means comprises:

means for obtaining conversion boundaries using the content value curves and scale factors for the modalities; and

means for converting the modality of the multimedia contents into the determined optimal modality.

10. The apparatus according to claim 9, wherein the conversion boundaries are values of the media resource corresponding to intersection points where the content value curves intersect with each other when the content value curves of the modalities overlap with each other according to the scale factors.

11. The apparatus according to claim 8, wherein each of the content value curves is obtained by combining content value curves that are measured according to two or more different qualities.

FIG. 1

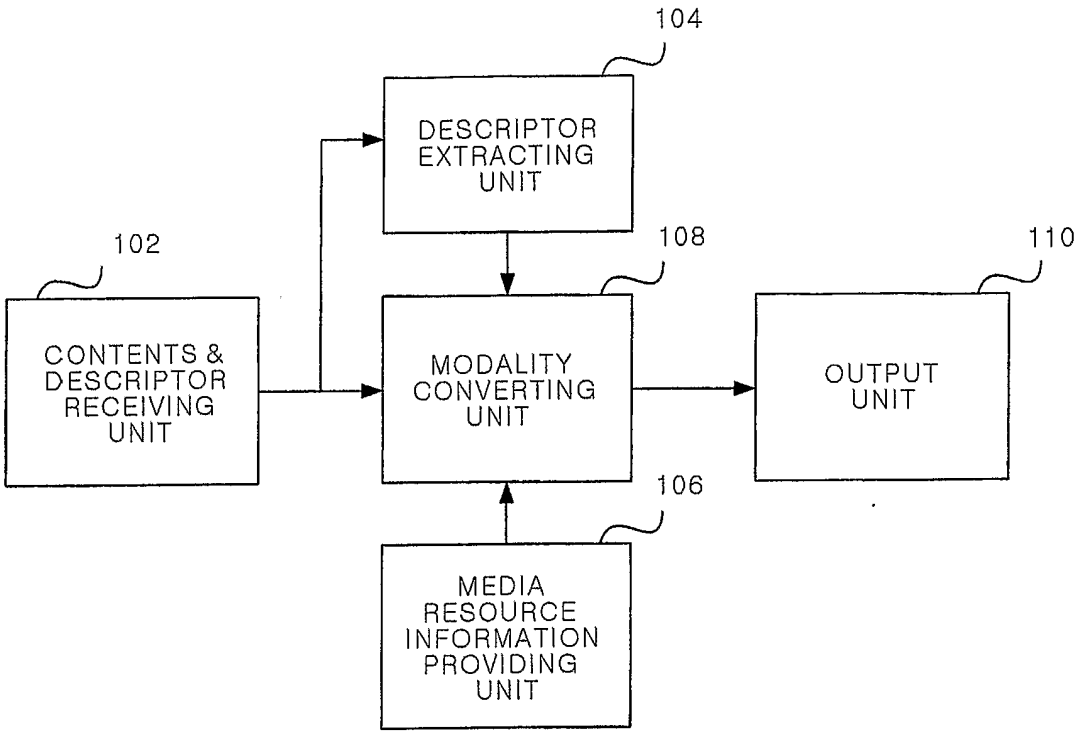

104
DESCRIPTOR EXTRACTING UNIT
102
108
110
CONTENTS & DESCRIPTOR RECEIVING UNIT
MODALITY CONVERTING UNIT
OUTPUT UNIT
106
MEDIA RESOURCE INFORMATION PROVIDING UNIT

FIG. 2

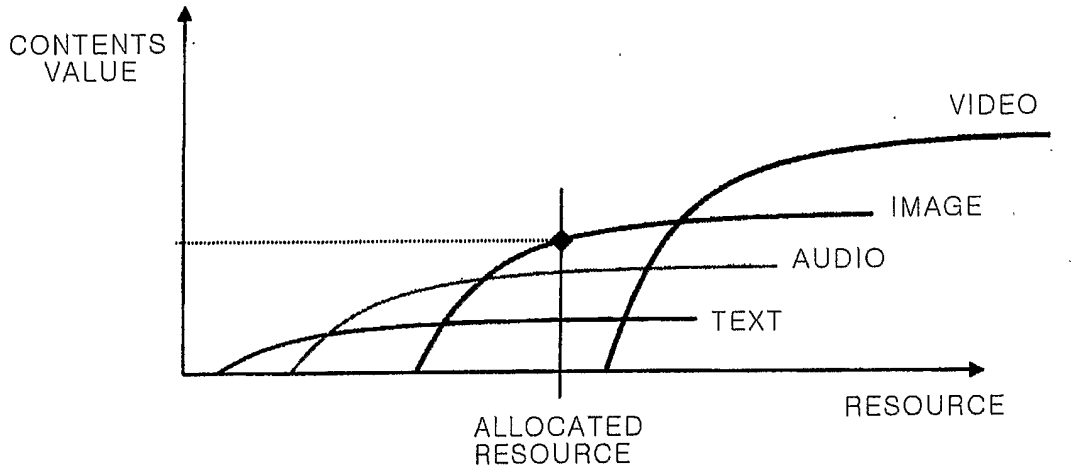

CONTENTS VALUE
VIDEO
IMAGE
AUDIO
TEXT
RESOURCE
ALLOCATED RESOURCE

FIG. 3

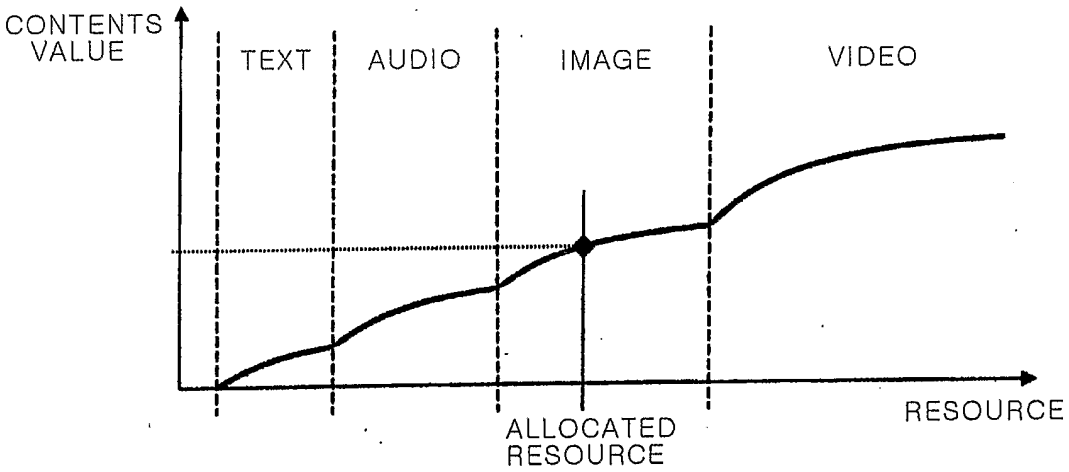

CONTENTS VALUE
TEXT
AUDIO
IMAGE
VIDEO
RESOURCE
ALLOCATED RESOURCE

FIG. 4a

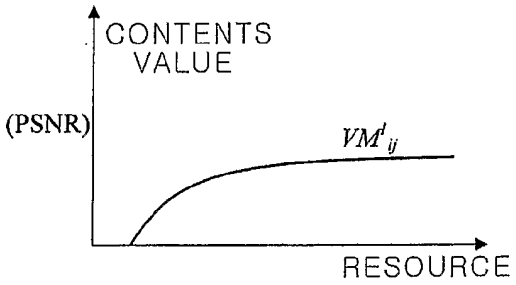
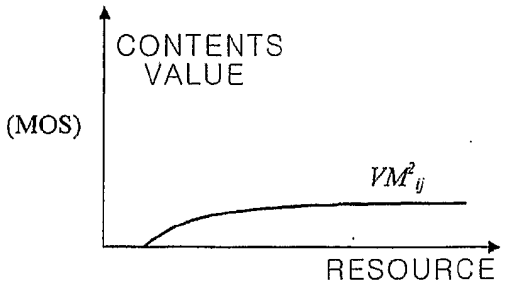
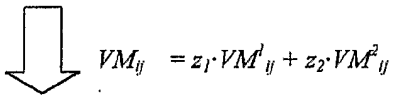


FIG. 4b





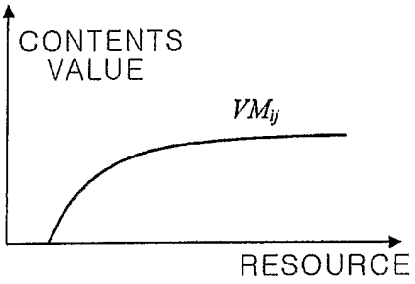
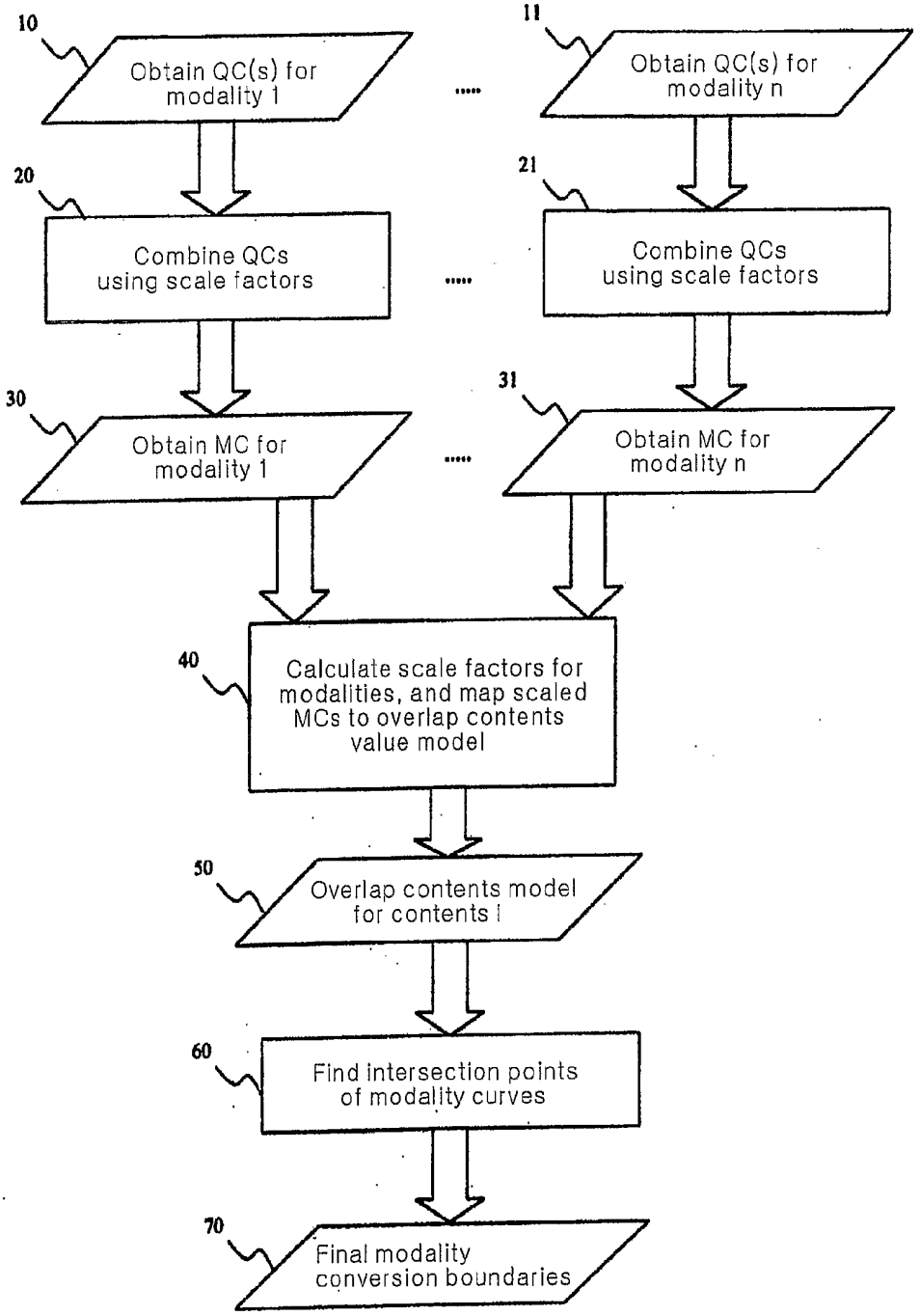


FIG. 4c

FIG. 5


10
Obtain QC(s) for modality 1
11
Obtain QC(s) for modality n
20
Combine QCs using scale factors
21
Combine QCs using scale factors
30
Obtain MC for modality 1
31
Obtain MC for modality n
40
Calculate scale factors for modalities, and map scaled MCs to overlap contents value model
50
Overlap contents model for contents i
60
Find intersection points of modality curves
70
Final modality conversion boundaries