TECHNICAL REPORT

ISO/IEC TR 15938-8

First edition
2002-12-15

AMENDMENT 2
2006-12-01

Information technology — Multimedia content description interface —

Part 8: Extraction and use of MPEG-7 descriptions

AMENDMENT 2: Extraction and use of MPEG-7 perceptual 3D shape descriptor

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC TR 15938-8:2002/Amd 2:2006
https://standards.iteh.ai/catalog/standards/sist/7bcdc5a3-fe8c-4615-8eef-a0a207a27c40/iso-iec-tr-15938-8-2002-amd-2-2006

Technologies de l'information — Interface de description du contenu multimédia —

Partie 8: Extraction et utilisation des descriptions MPEG-7

AMENDEMENT 2: Extraction et emploi du descripteur de forme 3D perceptuel MPEG-7

Reference number
ISO/IEC TR 15938-8:2002/Amd.2:2006(E)

© ISO/IEC 2006

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO/IEC TR 15938-8:2002/Amd 2:2006
https://standards.iteh.ai/catalog/standards/sist/7bada5c3-fe8c-4615-8acf-a0a207a27edd/iso-iec-tr-15938-8-2002-amd-2-2006

© ISO/IEC 2006

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

In exceptional circumstances, the joint technical committee may propose the publication of a Technical Report of one of the following types:

— type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;

— type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;

— type 3, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC TR 15938-8:2002/Amd 2:2006
https://standards.iteh.ai/catalog/standards/sist/7bada5c3-fe8c-4615-8acf-a0a207a27edd/iso-iec-tr-15938-8-2002-amd-2-2006

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Amendment 2 to ISO/IEC TR 15938-8:2002 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

NOTE This document preserves the sectioning of ISO/IEC TR 15938-8:2002 and its amendments. The text and figures given below are currently being considered as additions and/or modifications to those corresponding sections in ISO/IEC TR 15938-8:2002 and its amendments.

© ISO/IEC 2006 – All rights reserved

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO/IEC TR 15938-8:2002/Amd 2:2006
https://standards.iteh.ai/catalog/standards/sist/7bada5c3-fe8c-4615-8acf-a0a207a27edd/iso-iec-tr-15938-8-2002-amd-2-2006

Information technology — Multimedia content description interface —

Part 8: Extraction and use of MPEG-7 descriptions

AMENDMENT 2: Extraction and use of MPEG-7 perceptual 3D shape descriptor

Add after 2.2.2.49:

2.2.2.50 Attributed Relational Graph (ARG)

A graph whose nodes (vertices) and edges (links) contain unary attributes and dyadic attributes (describing the relation between the nodes), respectively. The graph is described in the form of a vector.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

2.2.2.51 Constrained Morphological Decomposition (CMD)

An algorithm, based on the mathematical concepts of morphology and convexity, to decompose a voxelized 3-D object into several parts.

ISO/IEC TR 15938-8:2002/Amd 2:2006
https://standards.iteh.ai/catalog/standards/sist/7bada5c3-fe8c-4615-8acf-a0a207a27edd/iso-iec-tr-15938-8-2002-amd-2-2006

2.2.2.52 Weighted Convexity (WC)

A volume-weighted sum of each part's convexity.

2.2.2.53 Weighted Convexity Difference (WCD)

A difference of two WCs before and after merging of two parts.

2.2.2.54 Initial Decomposition Stage (IDS)

The procedure of applying the CMD to a voxelized 3-D object, once.

2.2.2.55 Recursive Decomposition Stage (RDS)

The procedure of applying the CMD recursively to the result of the IDS or a previous RDS.

2.2.2.56 Iterative Merging Stage (IMS)

The procedure of merging parts in the result of the RDS iteratively using the WCD.

2.2.2.57 Earth Mover's Distance (EMD)

A kind of distance measure based on a solution [AMD2-2] to the transportation problem in graph theory.

© ISO/IEC 2006 – All rights reserved

2.2.2.58 Query by Example

A query to a content (e.g. image, 3D object, etc.) retrieval system whereby the information need is expressed visually, by providing an example of the kind of target content desired. This can be useful when the user has difficulty forming a query using key words or when text descriptions are not present in the database. For example if the user wants to find images of beaches, he/she can use any available image of a beach as the query and the retrieval system is expected to return images of beaches as results.

2.2.2.59 Query by Sketch

A query by example whereby the example content is a sketch, drawn by the user, reflecting the key visual attributes of the information need.

2.2.2.60 Query by Modified Example

A query by example whereby the example content is created by modifying an existing example (for example, using a graphical editing tool) so that it best expresses the information need.

Add after subclause 8.5:

8.6 Perceptual 3D shape

The Perceptual 3D Shape descriptor is a part-based representation of a 3D object expressed as a graph. In this context “node” is a vertex in the graph representation corresponding to a part in the 3D model. Such a representation facilitates object description consistent with human perception. The Perceptual 3D Shape descriptor supports ‘Query by example’. Furthermore, it provides unique functionalities, such as ‘Query by sketch’ and ‘Query by modified example’, which make the content-based retrieval system more interactive and intuitive in querying and retrieving similar 3D objects.

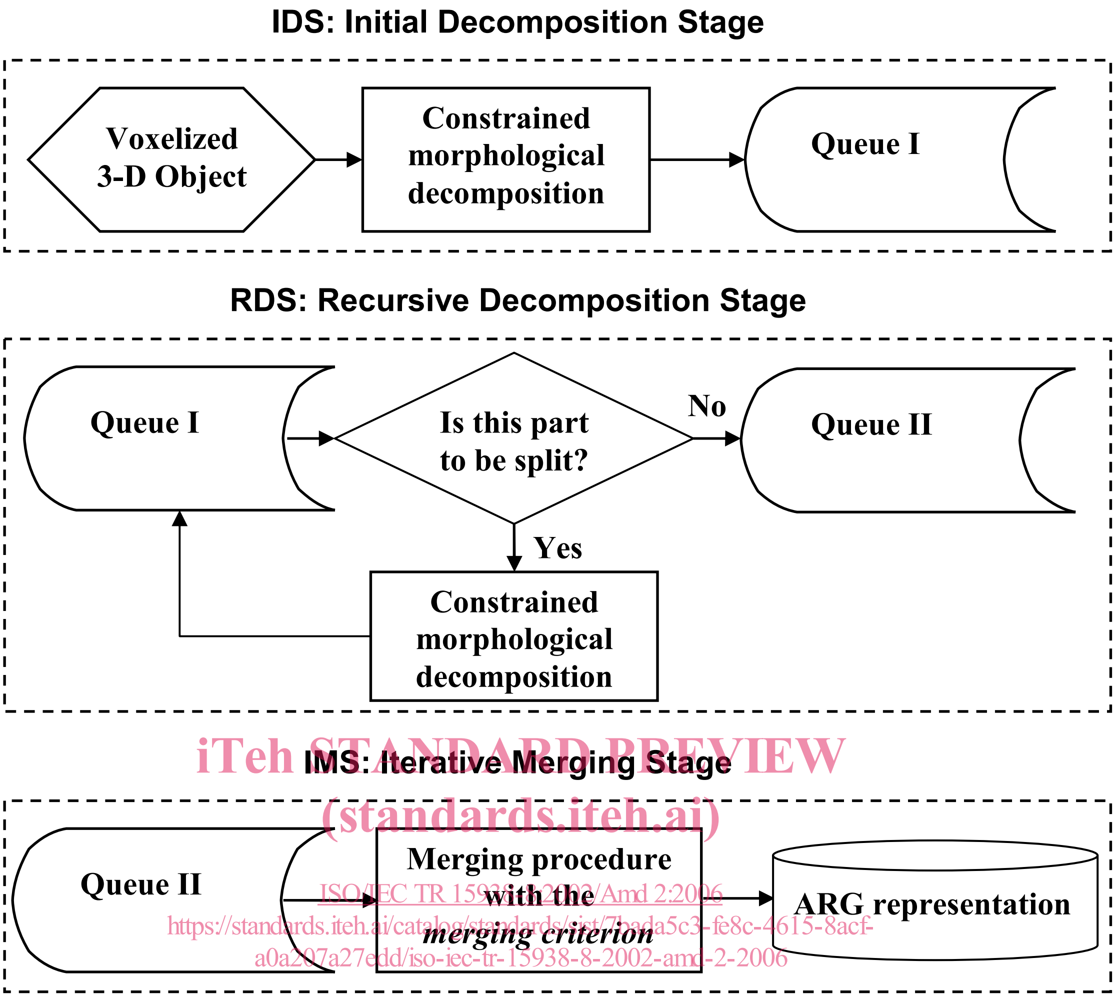
iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC TR 15938-8:2002/Amd 2:2006
https://standards.iteh.ai/catalog/standards/sist/7bade5c3-f68e-4615-8aef-
a0a2073a27eb/iso-iec-tr-15938-8-2002-amd-2-2006

8.6.1 Part-based representation

Part-based representation of 3D objects enables perceptual recognition that is robust in the presence of rotation, translation, deformation, deletion, and inhomogeneous scaling of a 3D object. More specifically, deletion and inhomogeneous scaling involve the removal of parts and growth or shrinkage of the specific part, respectively. In the task of forming a high-level object representation from low-level object features, parts serve as an intermediate representation.

The decomposition scheme [AMD2-1] is used to generate the attributed relational graph (ARG) of a 3D object. The proposed scheme recursively performs the constrained morphological decomposition (CMD) based on the mathematical morphology and weighted convexity. Then, a merging criterion based on the weighted convexity difference (WCD), which determines whether connected parts should be merged or not, is adopted for compact graph representation. The block diagram of the proposed scheme, in terms of three stages, is presented in Figure AMD2-1. The recursive decomposition stage (RDS) will be launched after the initial decomposition stage (IDS) and performed until QUEUE I is empty. Then, the iterative merging stage (IMS) is applied to parts in QUEUE II for the compact graph representation. Figure AMD2-2 shows the procedure of the proposed scheme for a ‘cow’ step by step. Figure AMD2-2 (a) and (b) show the ‘cow’ represented by rendered meshes and voxels, respectively. Then, Figure AMD2-2 (c), (d), and (e) show results of IDS, RDS, and IMS, respectively. Finally, the simple ARG representation is presented in Figure AMD2-2 (f), where the ellipsoidal node and edge represent the corresponding part and connectivity between parts, respectively.

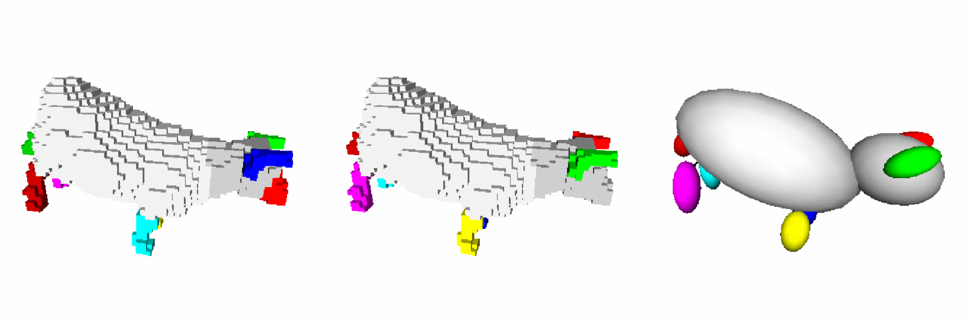
 © ISO/IEC 2006 – All rights reserved



iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC TR 15938-8:2002/Amd 2:2006
https://standards.iteh.ai/catalog/standards/sist/7bada5c3-fe8c-4615-8acf-a0a207a27edd/iso-iec-tr-15938-8-2002-amd-2-2006

Figure AMD2-1 — The block diagram of the decomposition scheme

(a) (b) (c)



(d) (e) (f)

Figure AMD2-2 — The procedure of generating a part-based representation

8.6.2 Feature extraction

As described in the previous subclause, the Perceptual 3D Shape descriptor has the form of an ARG, composed of nodes and edges. A node represents a meaningful part of the model with unary attributes, while an edge implies binary relations between nodes. In order to obtain all attributes, principal component analysis (PCA) is performed on every part of the 3D model to find three principle axes, where the 1st principal axis corresponds to the principal direction with biggest variance, and the 3rd axis corresponds to the direction with smallest variance. Afterwards,, 4 unary attributes and 3 binary relations are extracted to form a Perceptual 3D Shape descriptor. In detail, a node is parameterized by volume v, convexity c, and two eccentricity values e_1 and e_2. More specifically, the convexity is defined as the ratio of the volume in a node to that in its convex hull, and the eccentricity is composed of two coefficients, $e_1 = \sqrt{1 - c^2 / a^2}$ and $e_2 = \sqrt{1 - c^2 / b^2}$, where a, b, and c (a ≥ b ≥ c) are the maximum ranges along 1st, 2nd, and 3rd principal axes, respectively. Then edge attributes, i.e. binary relations between two nodes, are extracted from the geometric relation between two nodes, in which the distance between centers of connected nodes and two angles are adopted. The first angle is the angle between the 1st principal axes of the connected nodes and the other is between their 2nd principal axes. All the unary attributes and binary relations are normalized into the interval [0, 1]. However, to adopt 'Query by sketch' in the retrieval system, the Perceptual 3D Shape descriptor is required to be represented by the set of ellipsoids. In this context, each ellipsoid contains three properties, such as Volume, Max (i.e. maximum range along each principle axes) and Convexity, which can easily be converted into the 4 unary attributes. Next, the Perceptual 3D Shape descriptor contains three properties, such as Center, PCA_Axis_1 and PCA_Axis_2 (i.e. 1st and 2nd principle axis) from which the 3 binary relations can be computed. Therefore, an actual Perceptual 3D Shape descriptor is created, as shown in Binary Representation Syntax. Note that Volume, Center, Max and Convexity are in the interval [0, 1], while the components in PCA_Axis_1 and PCA_Axis_2 are in the interval [-1, 1].

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC TR 15938-8:2002/Amd 2:2006

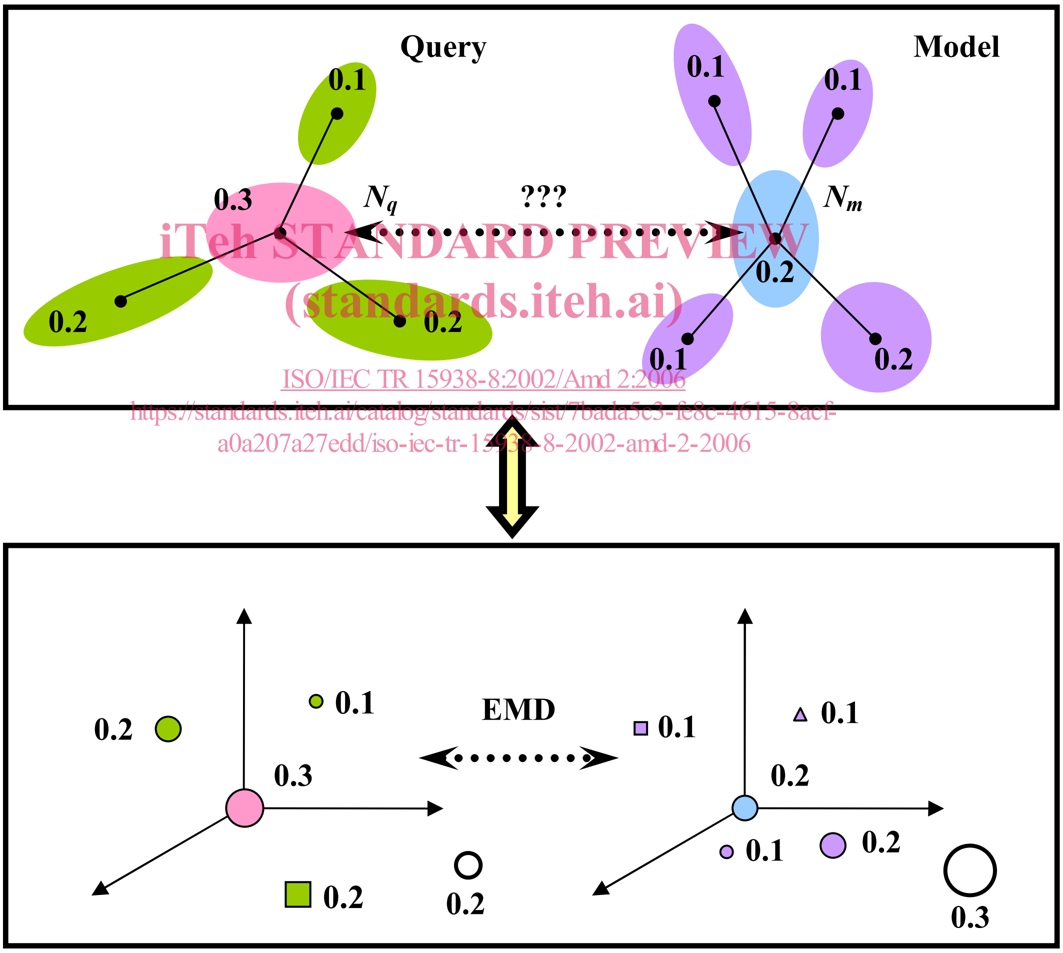
8.6.3 Similarity matching

The one-to-one comparison of two Perceptual 3D Shape descriptors consists of four steps: (1) Forming an ARG from every descriptor (Suppose that they are named as "query graph" and "model graph", respectively), (2) For each node in both graphs, defining Volume in Binary Representation Syntax as weight, (3) Calculating a distance matrix, where every element is the difference (distance) between any node-pair formed by any query graph node (named N_q) and any model graph node (named N_m), (4) Comparing the query and model graphs by employing the conventional Earth Mover's Distance (EMD) algorithm [AMD2-2], taking the node weights (from step-2) and the distance matrix (from step-3) as the input. In step-3 of this procedure, the calculation of the distance between N_q and N_m is fulfilled also by employing the EMD algorithm. This employment is named "Inner EMD", and the employment in step-4 is named "Outer EMD", thus the P3DS matching algorithm is named nested-EMD (nEMD).

Only step-3 needs more explanation. During this step, the distance between N_q and N_m is calculated as follows (from step-A to step-H): (A), their unary attributes are compared to give a "***unary-distance***". (B), a point set (named "Query Point Set") is constructed by N_q and all its connected nodes in the query graph. Every point is assigned a weight equal to the volume of its corresponding node. An imaginary point is also created and inserted to this point set, whose weight makes the sum of all set-points equal to one. (C) Another point set (named "Model Point Set") is constructed by N_m, all its connected nodes in the model graph, and a

 © ISO/IEC 2006 – All rights reserved

newly-introduced imaginary point, in the same way as step-B. (D) A vector space is constructed. The axes of its coordinate system represent the three measurements (one distance and two angles) of binary relations between connected nodes. N_q and N_m are located at the origin of this vector space, while other points (except imaginary points) are located so that their coordinate values are equal to their binary relation values with N_q (or N_m). Figure AMD2-3 shows the connected nodes and the point locations in the vector space respectively for the query and model point sets. (E) Two graphs (named "***Nq-Graph***" and "***Nm-Graph***", respectively) are constructed with the Query and Model Point Set, respectively. (F) A binary-distance matrix is constructed by calculating the distances between all node-pairs between Nq-Graph and Nm-Graph. For calculating every binary-distance matrix element, if neither node is an imaginary point, the Euclidean distance in the vector space is used (marked as dots in Figure AMD2-4); if one node is an imaginary point, a constant value ***d*** is used (marked as "d" in Figure AMD2-4); otherwise the distance is zero (marked as "0" in Figure AMD2-4). (G) The conventional EMD algorithm is employed (Inner EMD) for comparing Nq-Graph and Nm-Graph, taking the graph nodes from step-E and the binary-distance matrix from step-F as the input, to calculate a distance value (named "***second-distance***"). (H) The unary-distance from step-A and the second-distance from step-G are summed to give the final result.



iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC TR 15938-8:2002/Amd 2:2006
https://standards.iteh.ai/catalog/standards/sist/7bada5c3-fe8c-4615-8acf-a0a207a27edd/iso-iec-tr-15938-8-2002-amd-2-2006

Figure AMD2-3 — Vector space representation for computing the Inner EMD