

**Joint Video Experts Team (JVET)
of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC
29/WG 11**

Document: JVET-AL0147

MPEG: m71958

39th Meeting, Daejeon, KR, 26 June - 4 July 2025

Title: AHG16: Refined parameter translator of generative face video coding

Status: Input document to JVET

Purpose: Proposal

**Author(s) or
Contact(s):** [S. Yin](#) | shanzhyin3-c@my.cityu.edu.hk (CityUHK)
[Shurun Wang](#) | shurun.wsr@alibaba-inc.com (CityUHK)
[Zixiang Zhang](#) | zhangzixiang.zzx@alibaba-inc.com (CityUHK)
[Bolin Chen](#) | chenbolin.chenboli@alibaba-inc.com (Alibaba)
[Y. Ye](#) | yan.ye@alibaba-inc.com (Alibaba)
[Ru-Ling Liao](#) | ruling.lrl@alibaba-inc.com (Alibaba)
[Jie Chen](#) | jiechen.cj@alibaba-inc.com | submitter (Alibaba)

Source: CityUHK, Alibaba

HEADER EXCERPT FROM SOURCE DOCUMENT

Abstract

In JVET-AG0048, we investigated the issue of interoperability among different GFVC systems and solved the interoperability problem by adding face parameter translators to the AHG16 software. With these translators, it was shown that three different types of face parameters from different GFVC algorithms can be effectively adapted to one another while retaining promising coding performance. During 34th to 37th JVET meetings, several updates have been made to the GFVC models in the AHG16 software, including more algorithms, more diverse training data, multi-resolution models etc., which resulted in the existing parameter translators in the latest AHG16 software becoming outdated (e.g. inability to handle 512x512 resolution). In this contribution, we propose to update face parameter translators in the AHG16 software with a refined difference translation scheme. Experimental results show that refined translators can achieve translations between 4 multi-resolution models on both 256x256 and 512x512 resolutions while retaining promising coding performance.

Built from JVET-AL0147-v1.docx

Archive

DOC JVET-AL0147-v1.docx

XLS JVET-AL0147-v1.xlsm

```
graph TD; Archive[Archive] --> DOC[DOC JVET-AL0147-v1.docx]; Archive --> XLS[XLS JVET-AL0147-v1.xlsm];
```

The screenshot displays the Jvett Document Management System interface. At the top, a dark blue header contains the text "JVET DOCUMENT MANAGEMENT SYSTEM" and "Jvett Video Experts Team". Below the header, a light blue sidebar on the left lists navigation options: "User", "Create new user", "Claim your password", and "Contact". The main content area has a light blue background and features a document preview section for "JVET-AL0147". This section includes a red "Document Information" header, a "Submitted by" field with the name "Jana-Patrick Oden", a "Title" field with the text "FHC16: Refined parameter translator of generative face video coding", and an "Authors" field with the name "V. Sengul". To the right of this information is a section titled "Associated Resources". Below the document information is a "Document" section showing the file name "JVET-AL0147" and a link to "upload document". At the bottom, there is a "Comments" section with a red header and a large empty text area for user input. On the right side of the interface, there is a QR code.

  		
Joint Video Experts Team (JVET) of ITU-T SG21 WP3/21 and ISO/IEC JTC 1/SC 29 39th Meeting: Daejeon, KR, March 2025		Document: JVET-AL0147
Title:	Meeting Report of the 41st Meeting of the Joint Video Experts Team	
Status:	Report document from the chair of JVET	
Purpose:	Report	
1 Summary		
		

MEETING NOTES

[JVET-AL0147](#) AHG16: Refined parameter translator of generative face video coding [S. Yin, S. Wang, Z. Zhang (CityUHK), B. Chen, Y. Ye, R.-L. Liao, J. Chen (Alibaba)]

In JVET-AG0048, we investigated the issue of interoperability among different GFVC systems and solved the interoperability problem by adding face parameter translators to the AHG16 software. With these translators, it was shown that three different types of face parameters from different GFVC algorithms can be effectively adapted to one another while retaining promising coding performance. During 34th to 37th JVET meetings, several updates have been made to the GFVC models in the AHG16 software, including more algorithms, more diverse training data, multi-resolution models etc., which resulted in the existing parameter translators in the latest AHG16 software becoming outdated (e.g. inability to handle 512x512 resolution). In this contribution, we propose to update face parameter translators in the AHG16 software with a refined difference translation scheme. Experimental results show that refined translators can achieve translations between 4 multi-resolution models on both 256×256 and 512×512 resolutions while retaining promising coding performance.

The translation in feature difference domain provides results which are visually closer to the original than with previous translators between different GFVC algorithms. Also objective metrics are clearly improved.

Decision(SW): Add JVET-AL0147 to AHG branch of software, replace existing translation.

CITATION

S. Yin, S. Wang, Z. Zhang, B. Chen, Y. Ye, R.-L. Liao, J. Chen, "AHG16: Refined parameter translator of generative face video coding", JVET-AL0147, 39th Meeting: Daejeon, KR, March 2025.