

JVET-AE0088

AHG9: A study on Generative Face Video SEI Message

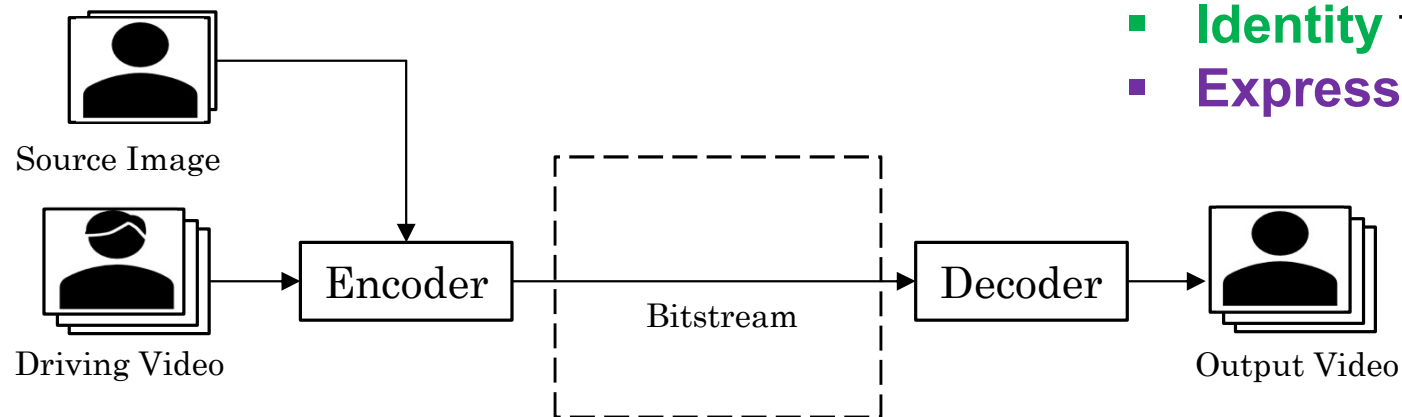
Han Boon Teo, Jing Yuan Thong, Karlekar Jayashree,
Chong Soon Lim, Kiyofumi Abe

Outline

- Motivation
- Compare 3D keypoints vs 3D landmarks
- 3D-Landmark-based Architecture
- Demo Videos
- Conclusion

Motivation

Face Re-enactment

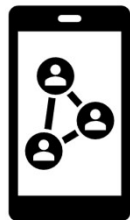


- **Identity** from **Source** Image
- **Expressions, Pose** from **Driving** Video

Use Cases



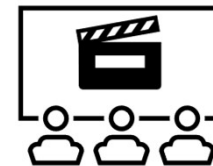
Video Conferencing



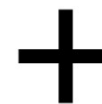
Social Media



Video Games



Entertainment



And more...

3D keypoints vs 3D landmarks

Keypoints

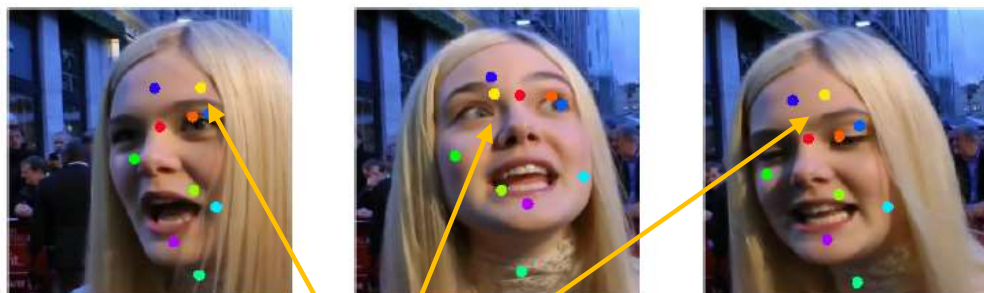
vs

Landmarks

Frm N

Frm N+1

Frm N+...

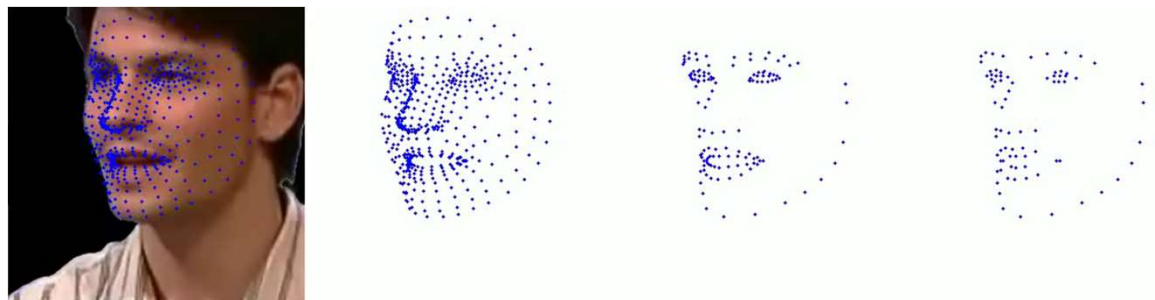


Within same sequence, keypoints not fixed at same location of face or frame

478 points

128 points

78 points



Within same sequence, landmarks linked to same feature point on face

3D keypoints vs 3D landmarks

Keypoints

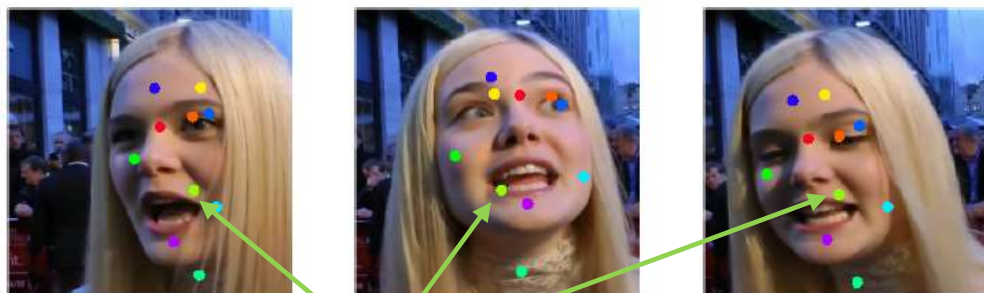
vs

Landmarks

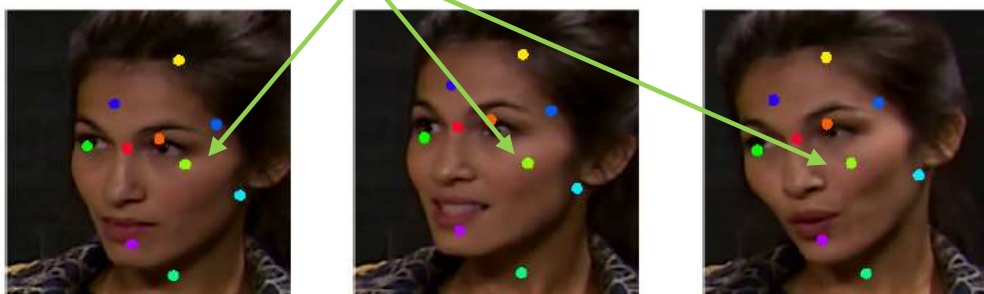
Frm N

Frm N+1

Frm N+...



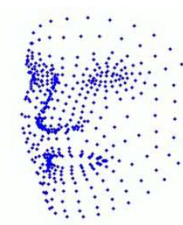
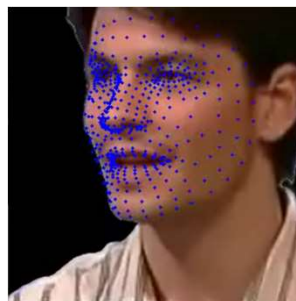
Across different sequences, keypoints not fixed at same location of face



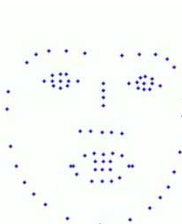
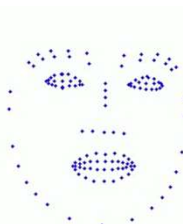
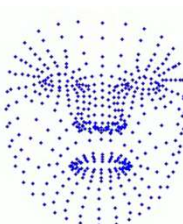
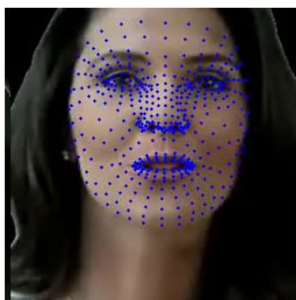
478 points

128 points

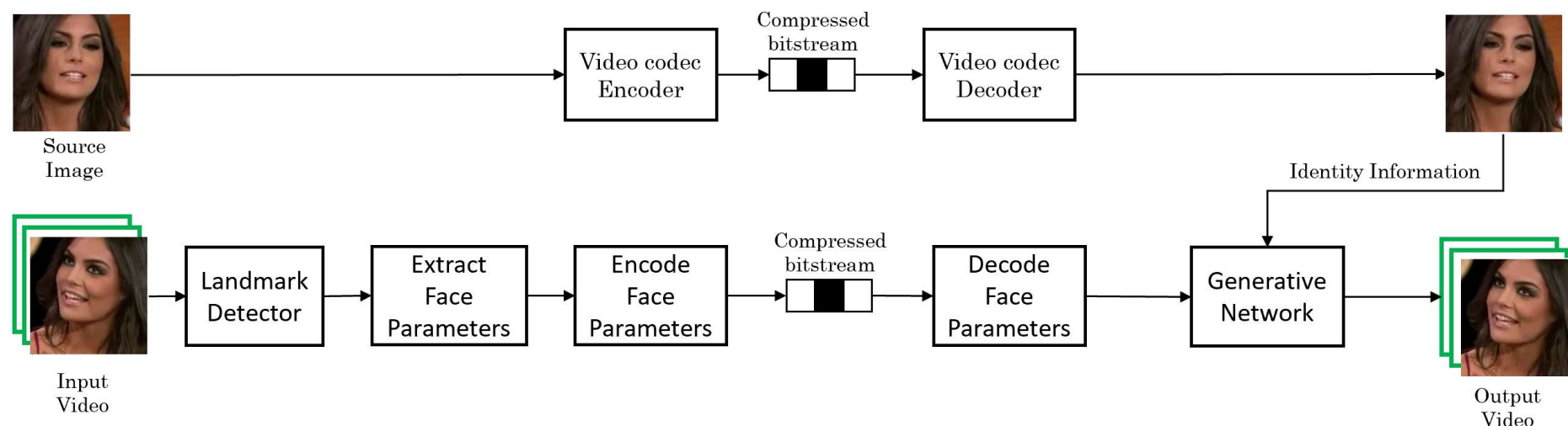
78 points



Across different sequences, landmarks linked to same feature point on face



3D-Landmark-based Face Re-enactment Architecture

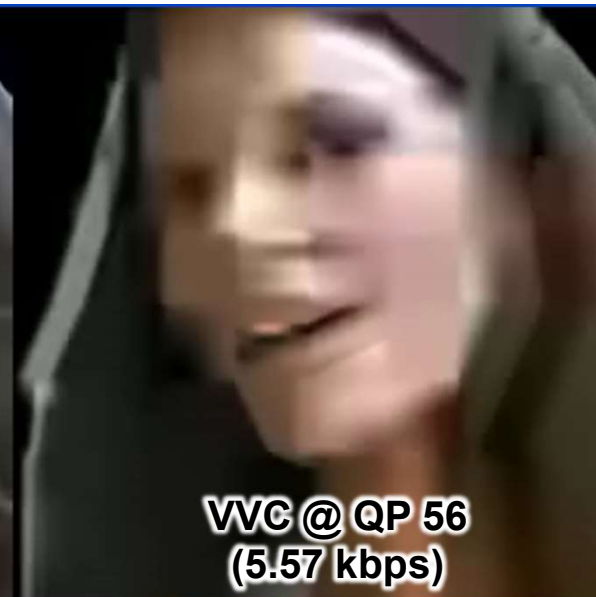


Highlights

- ✓ Integrated solution with video codec (VVC, HEVC or AVC)
- ✓ Landmarks coded in SEI messages of video bitstream
- ✓ Supported ultra-low bandwidth

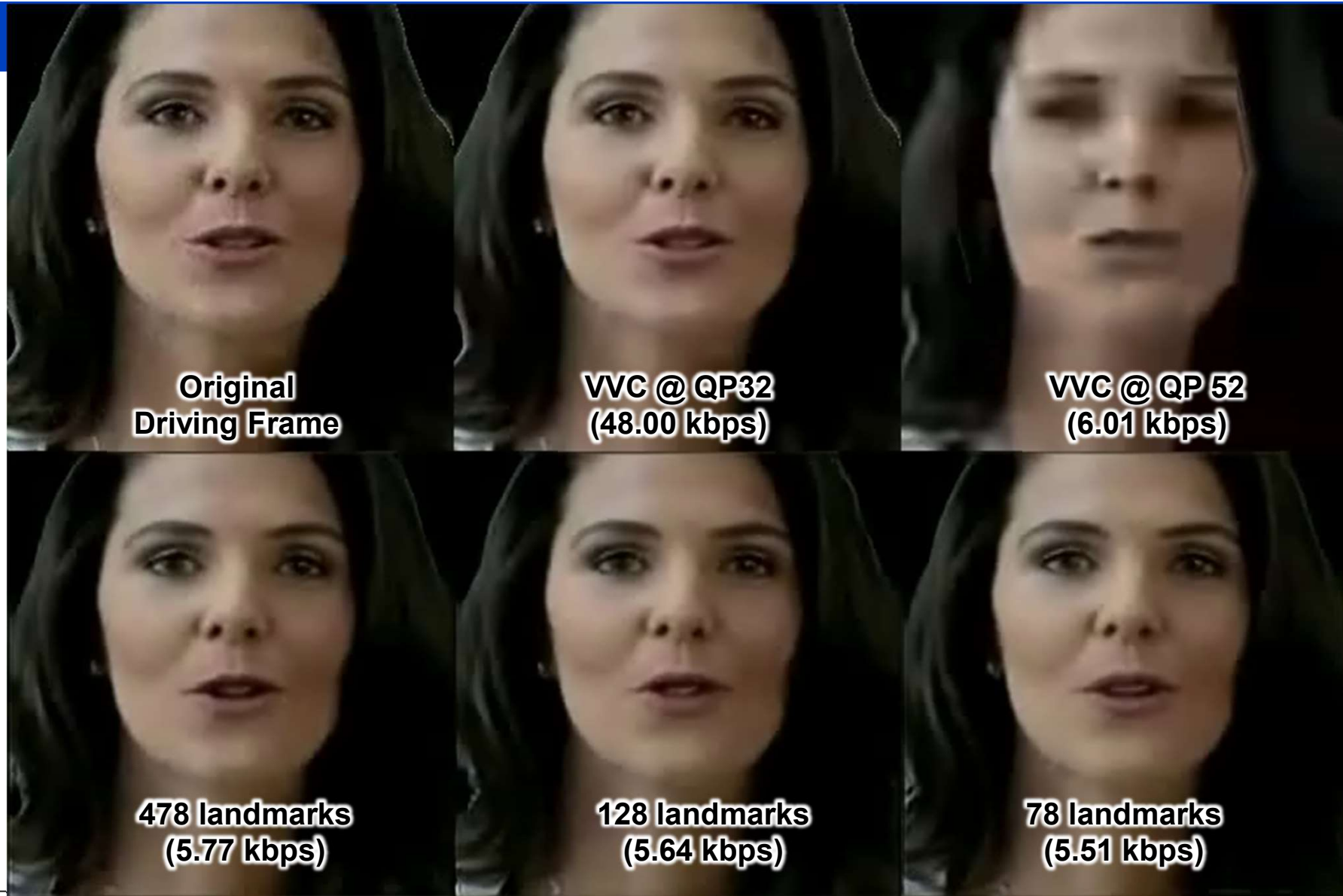
Demo

- Similar visual quality with VVC @ QP32
- Better visual quality than VVC at similar bitrate



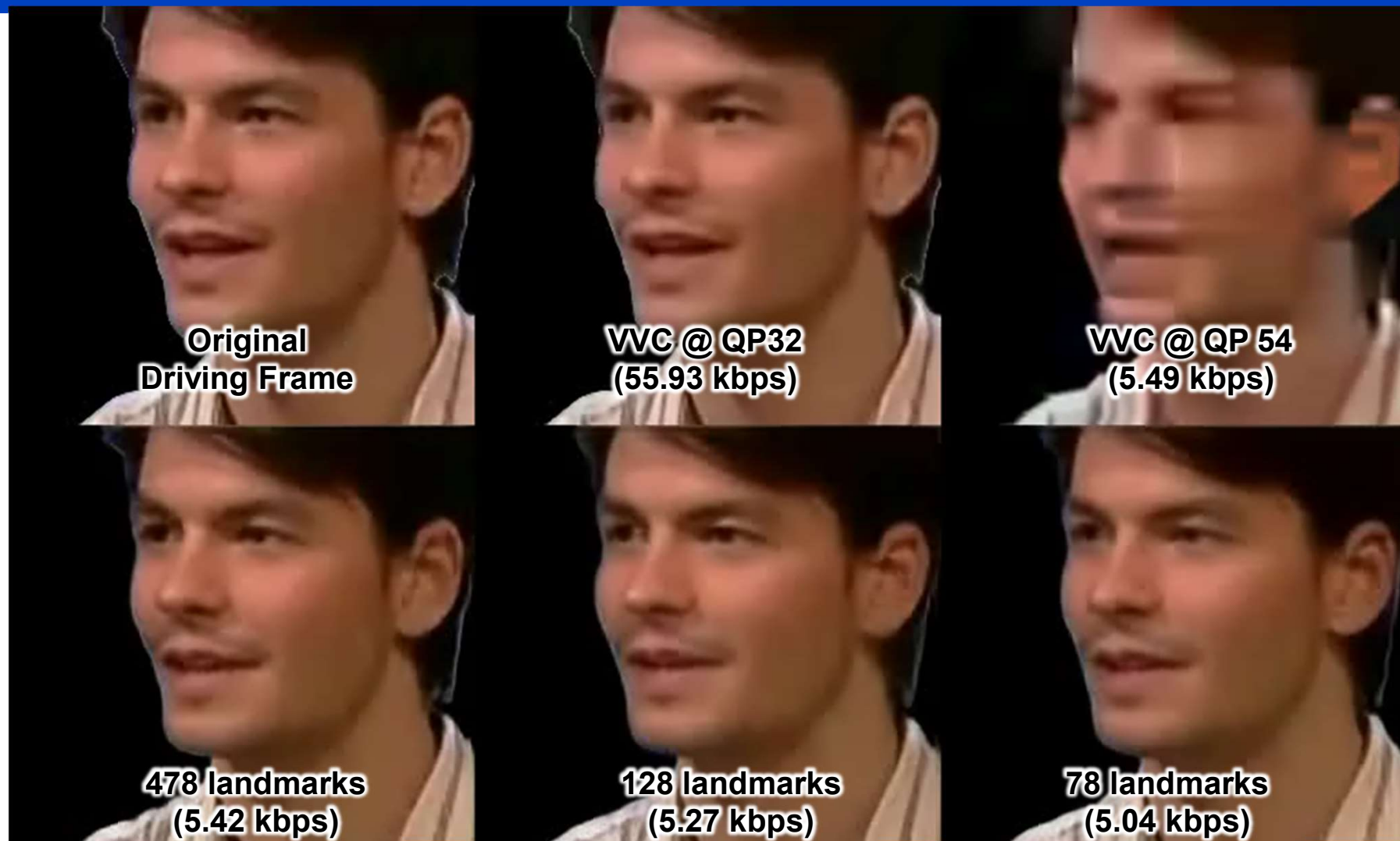
Demo

- Similar visual quality with VVC @ QP32
- Better visual quality than VVC at similar bitrate



Demo

- Similar visual quality with VVC @ QP32
- Better visual quality than VVC at similar bitrate



Conclusion

- ✓ 3-D landmarks should be considered for face re-enactment
- ✓ Advantages
 - ❑ **Interpretable** representations
 - ❑ **Flexible** against varying number of landmarks
 - ❑ Cater for different representations at encoders and decoders
 - **Compatible across different devices**
 - ❑ Can change landmark detector in future iterations
 - **Backward-compatibility**