JEG-Hybrid large database approach

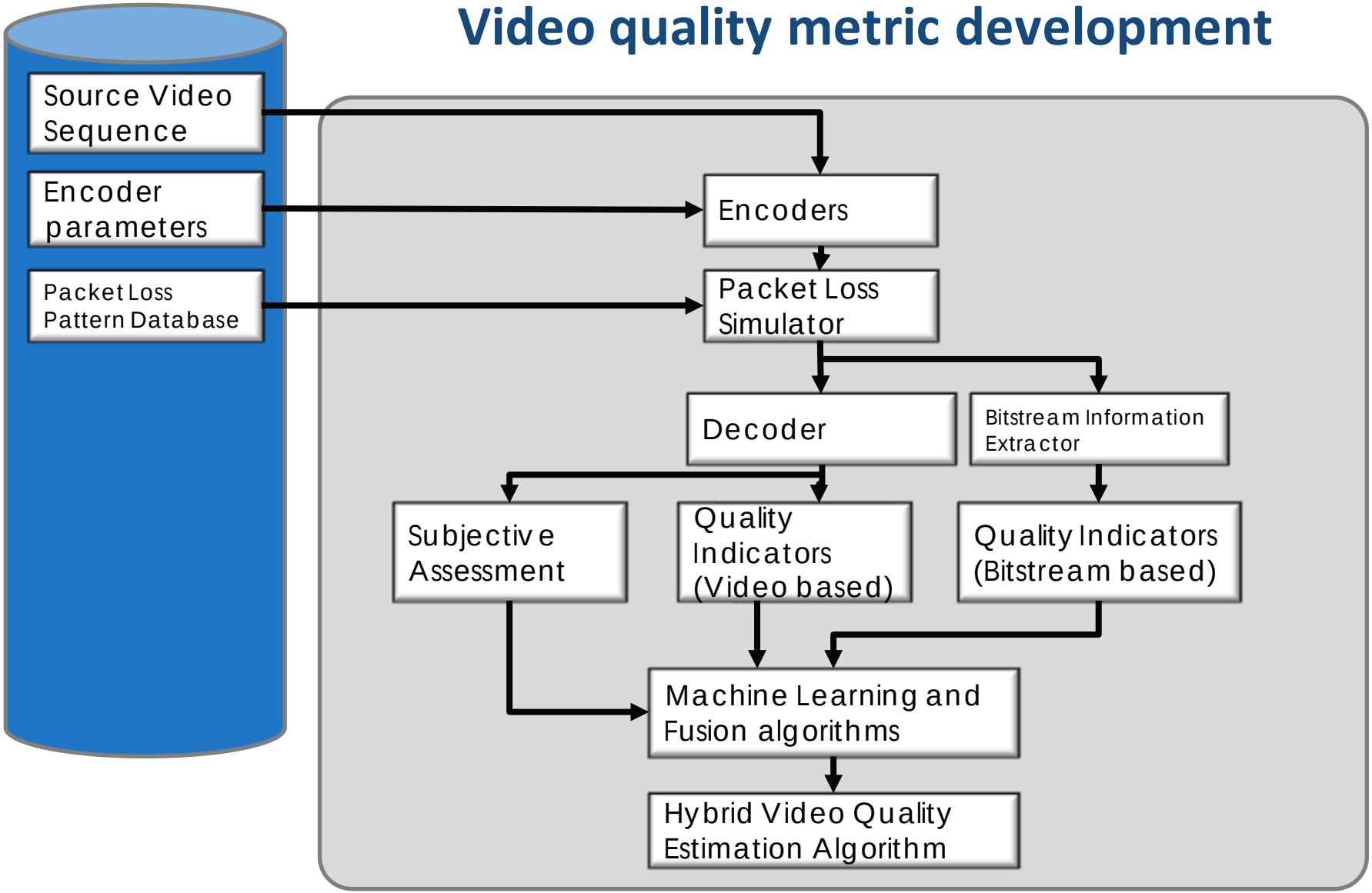
overview

Video quality metric development

- Key aspect is subjective evaluation
- Difficulties:
  - Assemble a representative and large enough group of people.
  - Controlled environment (BT.500) or equipment to perform the test in normal environment.
  - The time people concentrate on the job is limited. Therefore, in general restrict to 45min/person.

- Conclusion: subjective evaluation is the limiting aspect in video quality metric development.

Video quality metric development



Video quality metric development

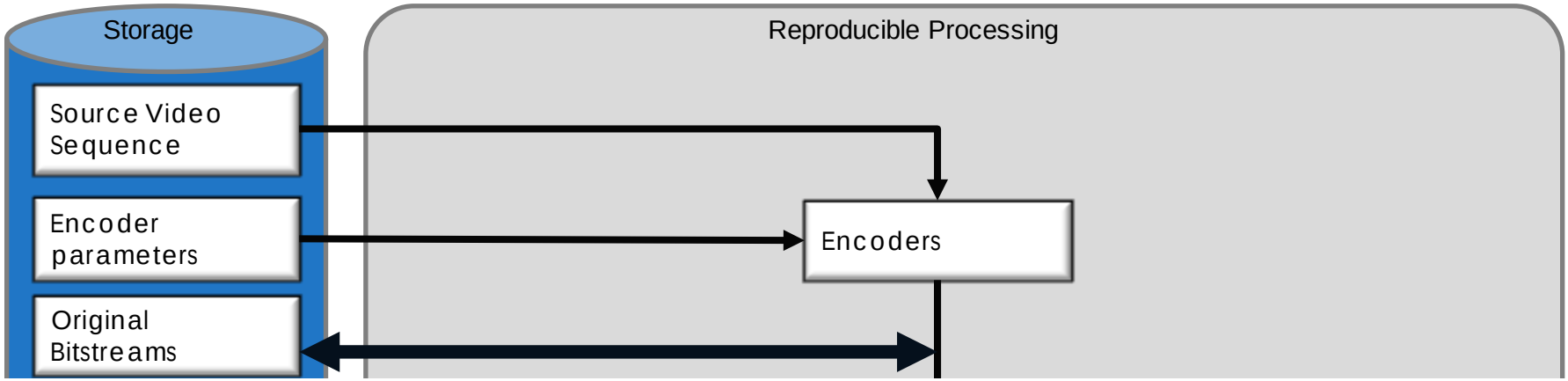
- Limited amount of source sequences:
  - consider broad SI-TI space
- Limited amount of compression settings:
  - Intra period
  - Bitrate
  - Slices
  - GOP

  } realistically chosen, but limited
- Limited amount of packet loss scenarios
  - Based on Packet Loss Rate percentages

- General applicability of quality metric can not be assumed.

Proposed large database approach

- By using
  - a large database covering
    - as much as possible sequences
    - as much as possible relevant encoder settings
    - as much as possible relevant packet loss scenarios
  - combined with the information delivered from existing full-reference quality metrics
  - and an as limited as possible subjective evaluation
- we try to design a general applicable video quality metric.


Storage
Source Video Sequence
Encoder parameters
Original Bitstreams
Reproducible Processing
Encoders


10 HD
31 UHD
Türkiye Joker
NERC DTV

HD

HEVC
HM 11

Bitrate	500kbps, 1Mbps, 2Mbps, 4Mbps, 8Mbps, 16Mbps
Rate control	Picture adaptive, CU adaptive
QP	26, 32, 38, 46
GOP	1, 2, 4, 8
Intra period	8, 16, 32, 64
Random access	IDR (closed-GOP intra), CDR (open-GOP intra)
Resolution	1920x1080, 1280x720, 960x544
Slices	0, 2, 4, 1500byte

AVC
Reduced
JM
x264

QP	24, 28, 32, 36, 40, 44
GOP	8
Intra period	64
Random access	CDR (open-GOP intra)
Resolution	3840x2160, 1920x1080, 1280x720
Slices	0
Min/Max block size	16/16, 8/16, 16/32, 8/32, 16/64, 8/64
Search range	16, 32, 64, 96, 128, all

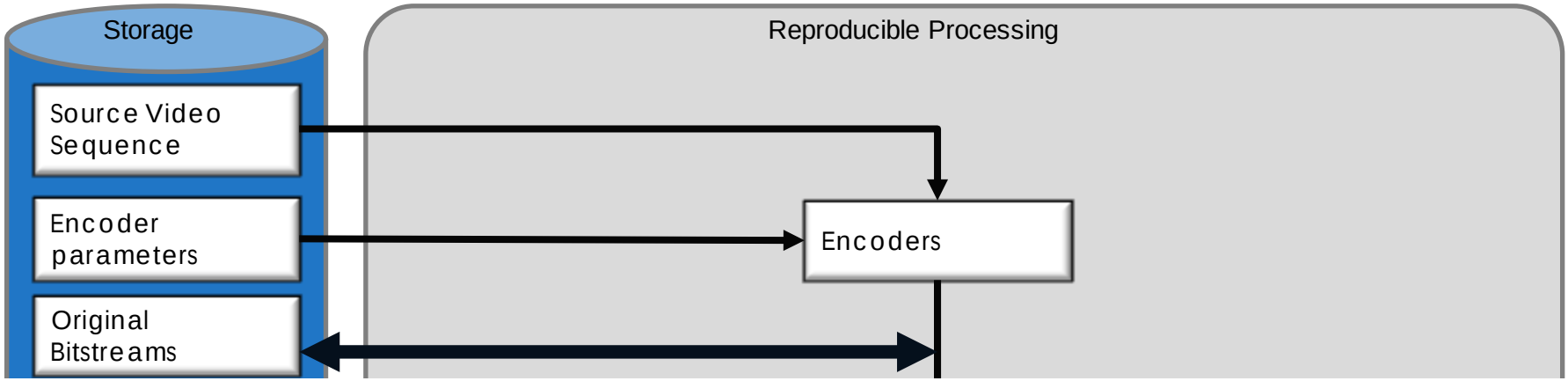
UHD

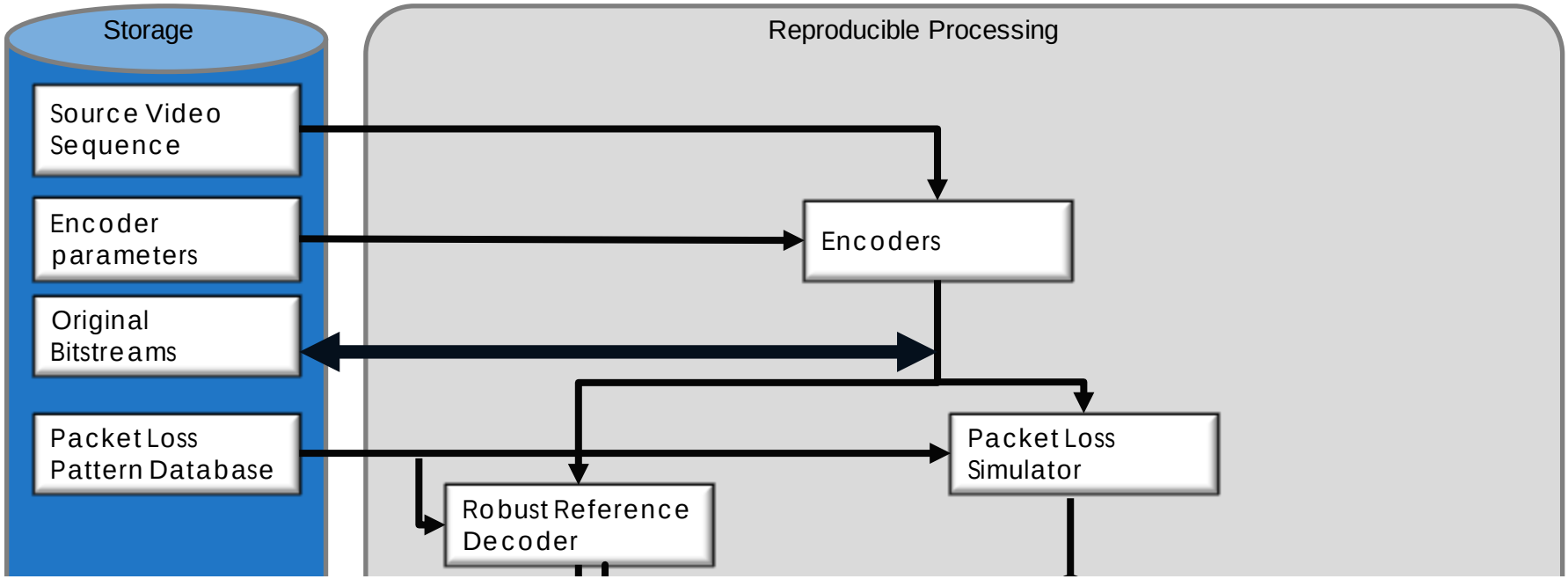
HM 14

- HD AVC
- 13000 PVS

- HD HEVC
- 60000 PVS

- UHD HEVC
- 40000 PVS


Storage
Source Video Sequence
Encoder parameters
Original Bitstreams
Reproducible Processing
Encoders


Storage
Source Video Sequence
Encoder parameters
Original Bitstreams
Packet Loss Pattern Database
Reproducible Processing
Encoders
Packet Loss Simulator
Robust Reference Decoder

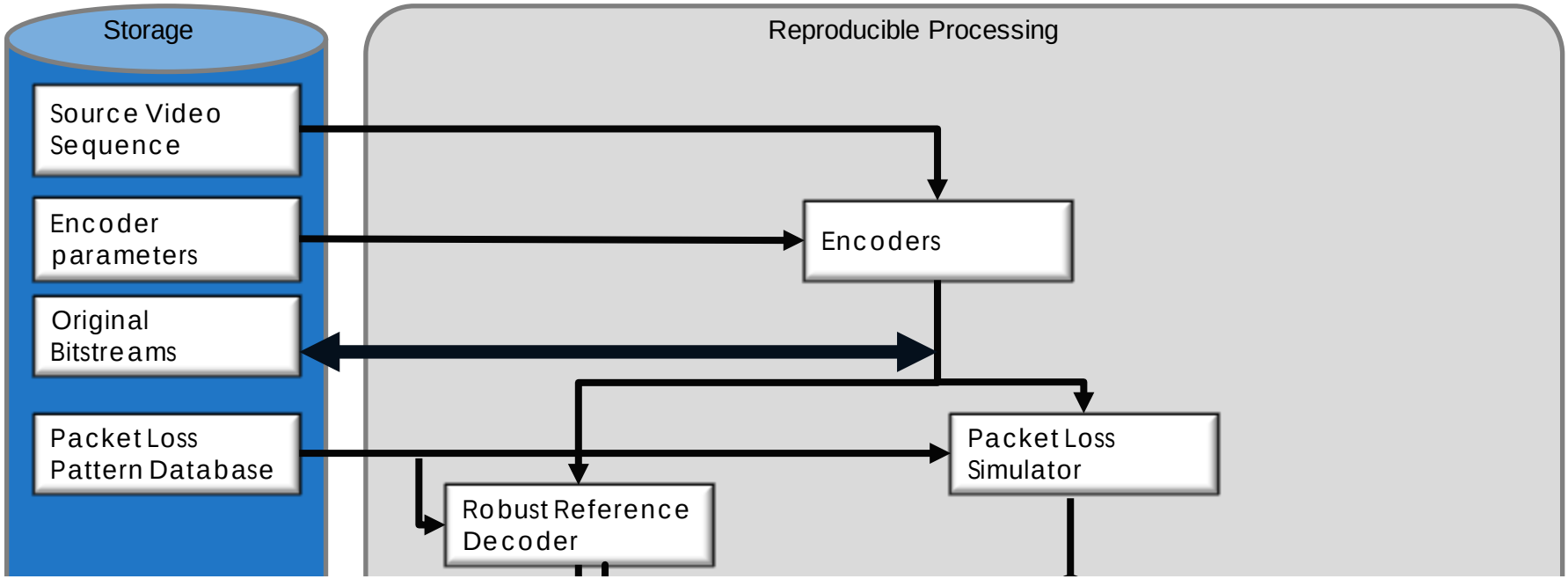
Two packet loss introducing tools

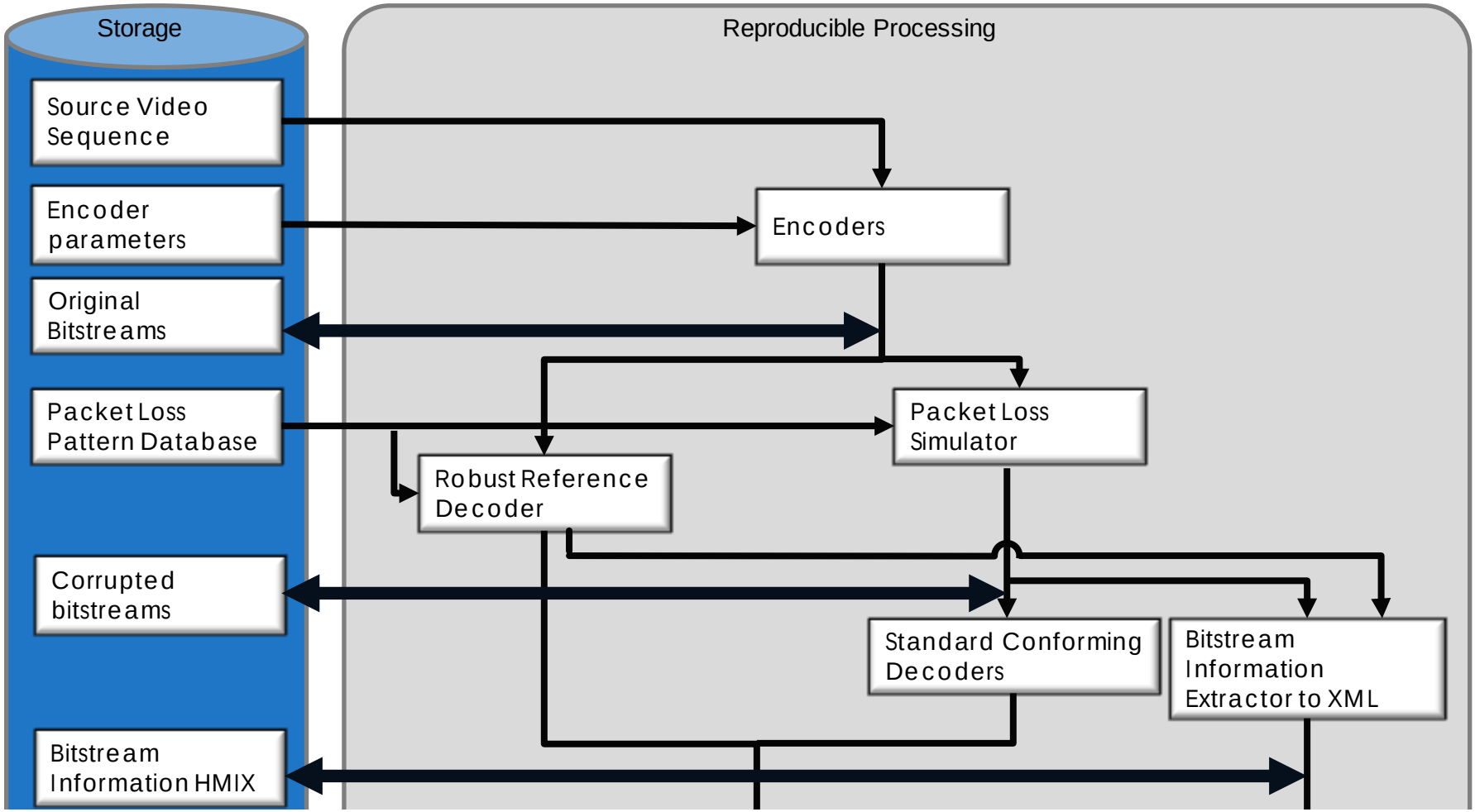
- Packet loss simulation – Sirannon
  - Video stream is encapsulated in RTP or MPEG-TS
  - Network transmission is simulated
  - Network packets are removed

  Downside is non-compliance of video stream, so decodability is not guaranteed

- Robust reference decoder
  - Decoder handles frame loss by reference buffer management

  Advantage is that every frame loss scenario results in decodability

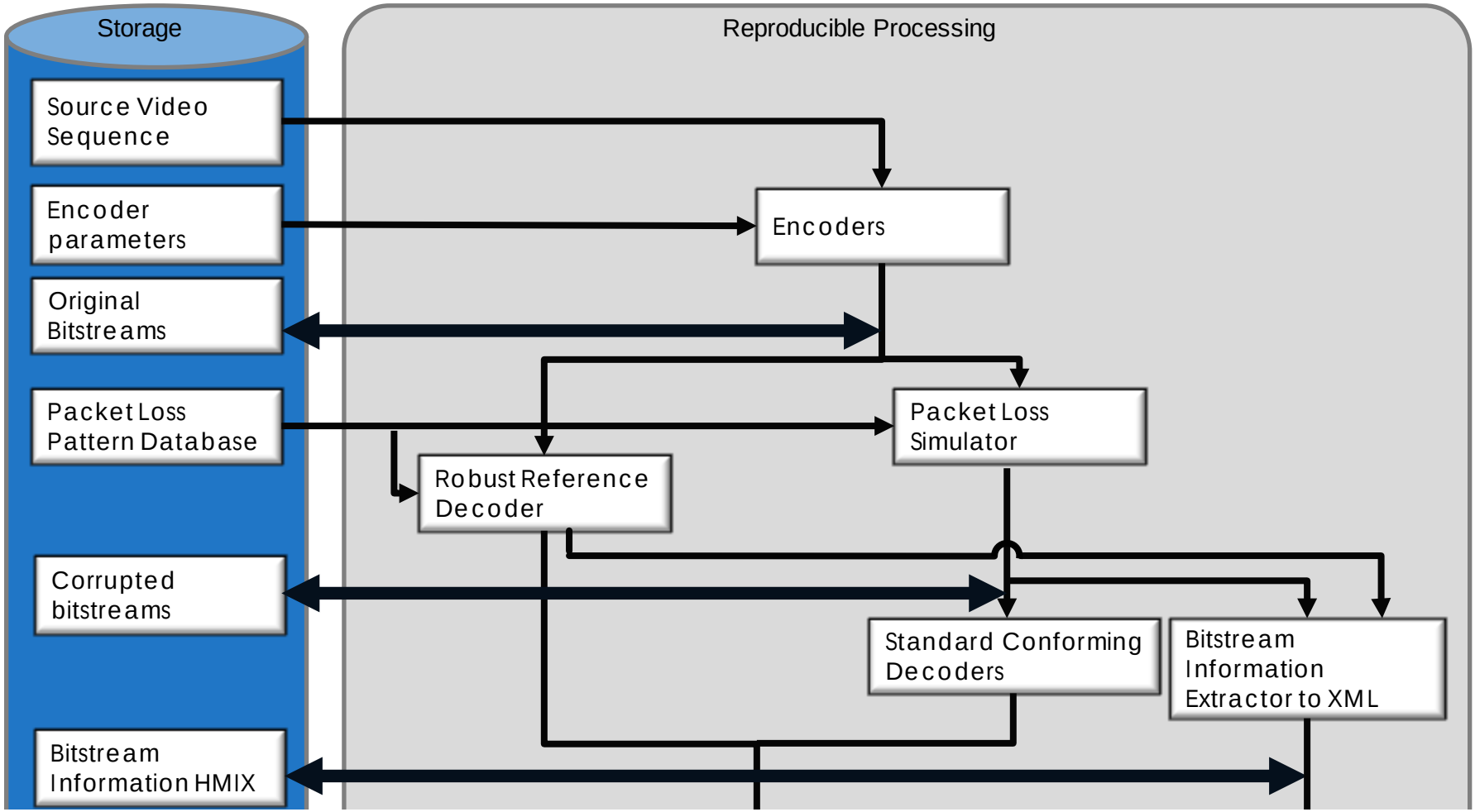

Storage
Source Video Sequence
Encoder parameters
Original Bitstreams
Packet Loss Pattern Database
Reproducible Processing
Encoders
Packet Loss Simulator
Robust Reference Decoder

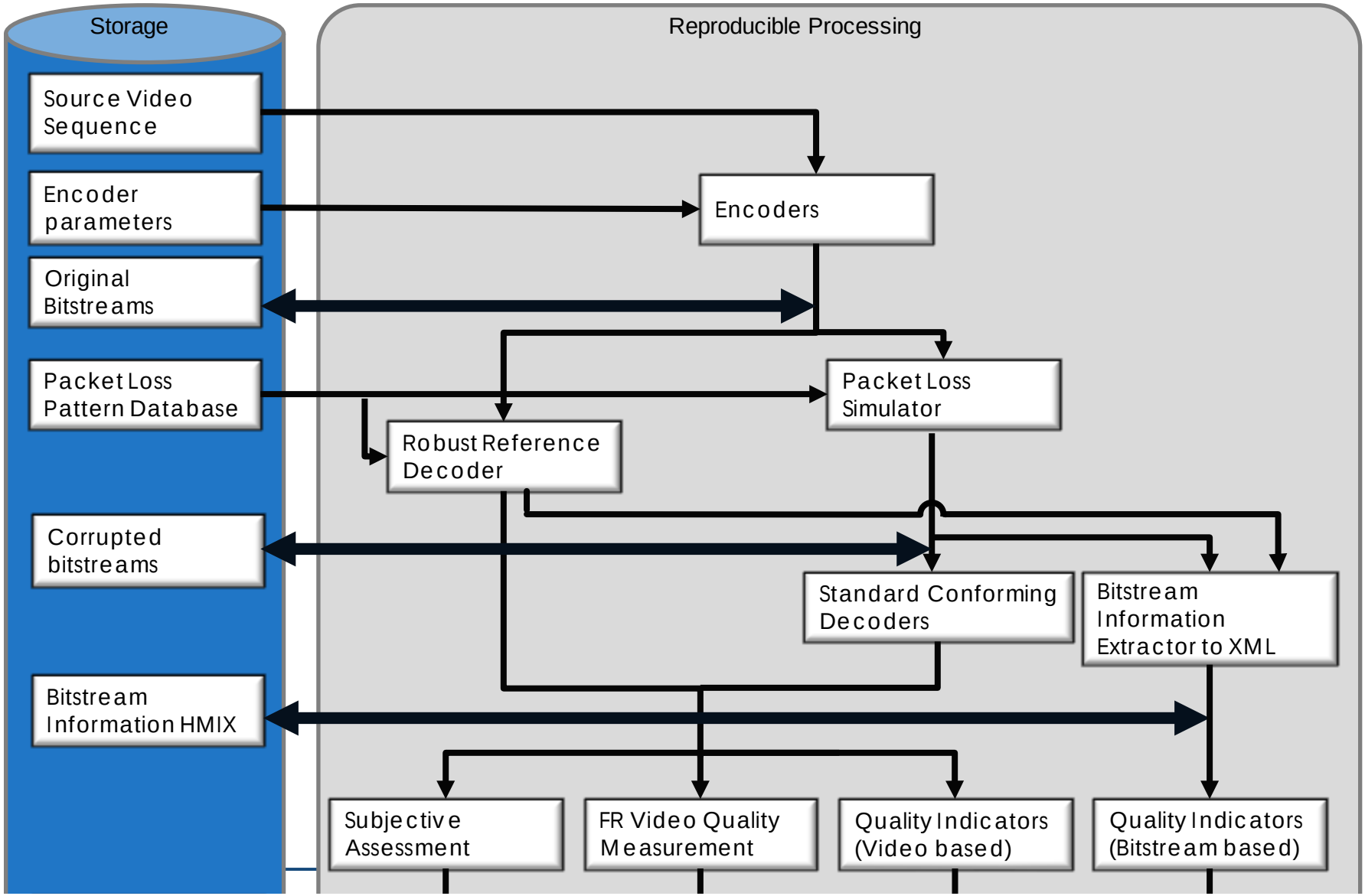

Storage
Reproducible Processing
Source Video Sequence
Encoder parameters
Original Bitstreams
Packet Loss Pattern Database
Corrupted bitstreams
Bitstream Information HMIX
Encoders
Packet Loss Simulator
Robust Reference Decoder
Standard Conforming Decoders
Bitstream Information Extractor to XML

HEVC video stream information extraction

- Machine readable (XML) formatted information from the video stream:
  - Block size
  - Prediction mode (intra/inter)
  - Quantization information
  - Motion vector information
  - Reference frame information

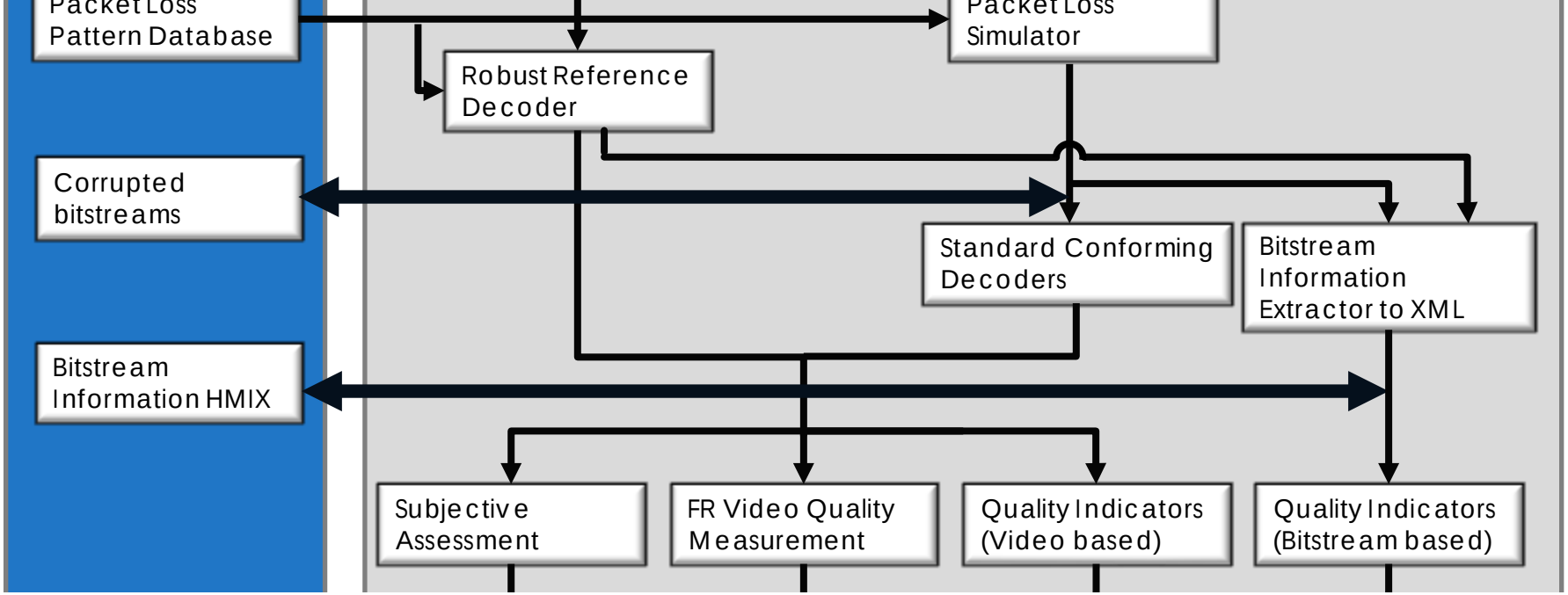
- This tool provides video stream information, so it is applicable in other domains as well:
  - Compressed domain object detection
  - Video stream transcoding

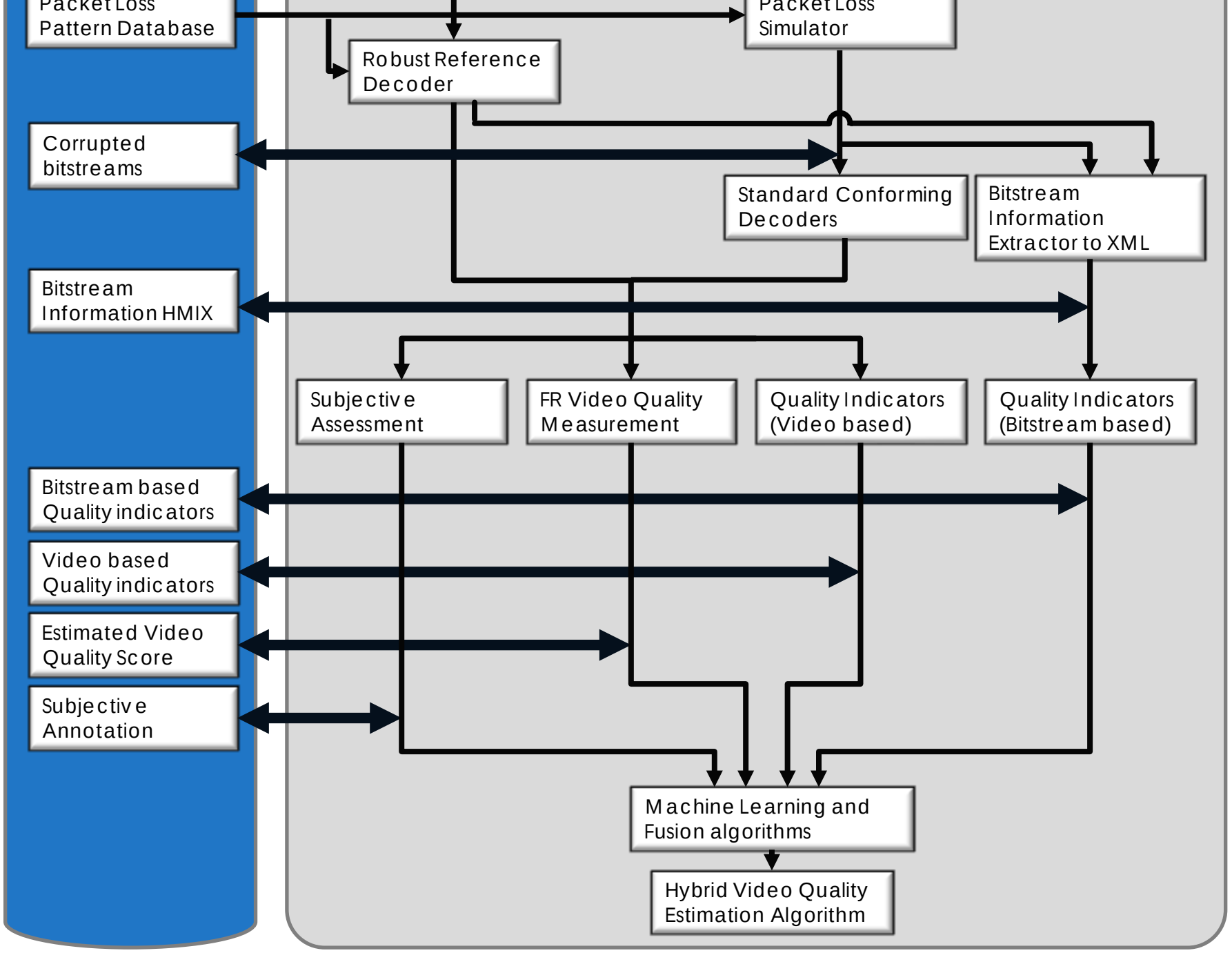

Storage
Reproducible Processing
Source Video Sequence
Encoder parameters
Original Bitstreams
Packet Loss Pattern Database
Corrupted bitstreams
Bitstream Information HMIX
Encoders
Packet Loss Simulator
Robust Reference Decoder
Standard Conforming Decoders
Bitstream Information Extractor to XML


Storage
Reproducible Processing
Source Video Sequence
Encoder parameters
Original Bitstreams
Packet Loss Pattern Database
Corrupted bitstreams
Bitstream Information HMIX
Encoders
Packet Loss Simulator
Robust Reference Decoder
Standard Conforming Decoders
Bitstream Information Extractor to XML
Subjective Assessment
FR Video Quality Measurement
Quality Indicators (Video based)
Quality Indicators (Bitstream based)

Full-reference quality indicators

- PSNR
- SSIM
- VIF
- VQM
- PVQM


Packet Loss Pattern Database
Packet Loss Simulator
Robust Reference Decoder
Corrupted bitstreams
Standard Conforming Decoders
Bitstream Information Extractor to XML
Bitstream Information HMIX
Subjective Assessment
FR Video Quality Measurement
Quality Indicators (Video based)
Quality Indicators (Bitstream based)


Packet Loss Pattern Database
Packet Loss Simulator
Robust Reference Decoder
Corrupted bitstreams
Standard Conforming Decoders
Bitstream Information Extractor to XML
Bitstream Information HMIX
Subjective Assessment
FR Video Quality Measurement
Quality Indicators (Video based)
Quality Indicators (Bitstream based)
Bitstream based Quality indicators
Video based Quality indicators
Estimated Video Quality Score
Subjective Annotation
Machine Learning and Fusion algorithms
Hybrid Video Quality Estimation Algorithm

Observations

Table 2. Correlation between different measurements averaged over the entire sequence.

		Pearson	Spearman
PSNR	SSIM	0.52	0.77
VIF	SSIM	0.93	0.99
PSNR	VIF	0.61	0.81

Table 3. Correlation between different measurements averaged over the entire sequence excluding src09.

		Pearson	Spearman
PSNR	SSIM	0.84	0.97
VIF	SSIM	0.93	0.99
PSNR	VIF	0.94	0.97

- PSNR can be misleading:
  - value grows to infinity when encoded lossless.
  - happens on black frames (src09)

Observations

- At high quality, most metrics agree.
- At the lower quality end, disagreement starts to appear.

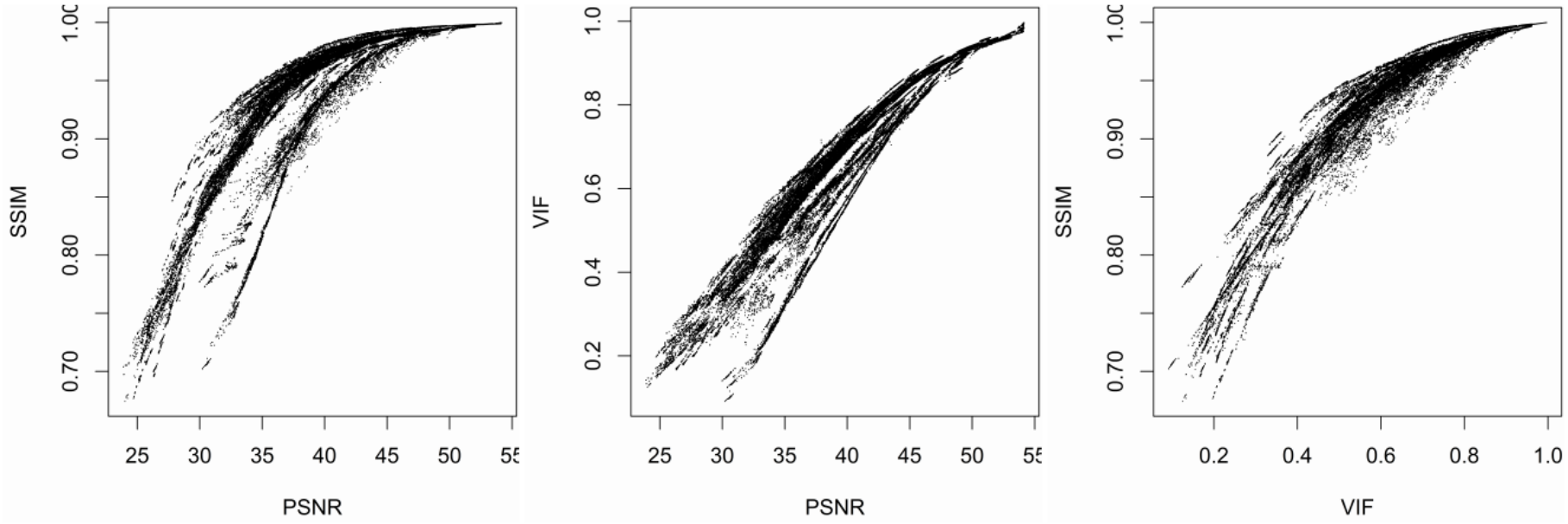


Fig. 3. Comparison between the different evaluated full-reference measurements.

Observations

- Analyze pairs of sequences: (i.e., 1,771,285,440 pairs):
- Disagreement between metrics in about 10.5% of the cases.
  - Due to PSNR: 55%
  - Due to SSIM: 30%
  - Due to VIF: 15%
- Sequence dependent

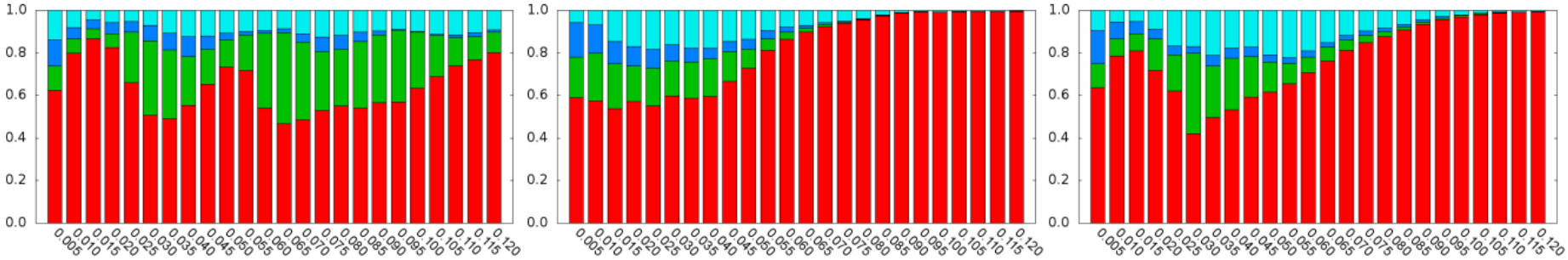
Table 4. Reasons of disagreement among quality measurements for each sequence.

Sequence	Pairs with disagreement	Due to PSNR	Due to SSIM	Due to VIF
src01	3.32%	14.47%	60.72%	24.80%
src02	2.64%	40.74%	45.70%	13.56%
src03	6.27%	61.97%	9.30%	28.73%
src04	4.55%	51.17%	11.76%	37.06%
src05	3.30%	37.89%	18.16%	43.95%
src06	4.99%	28.92%	13.84%	57.24%
src07	6.17%	69.45%	7.41%	23.14%
src08	3.93%	24.58%	59.33%	16.09%
src09	7.65%	20.89%	53.62%	25.49%
src10	3.81%	39.76%	12.55%	47.70%

Observations

- Normalized distance

$$\hat{d} = \sqrt{\Delta\widehat{PSNR}^2 + \Delta\widehat{SSIM}^2 + \Delta\widehat{VIF}^2}$$



(src03) (src05) (src10)

Fig. 4. Reason of disagreement (expressed as a ratio over the total pairs) between the various algorithms as a function of the normalized difference for some sequences, shown in brackets (red: agreement, green: due to PSNR, blue: due to SSIM, light blue: due to VIF).

- With a bigger difference in quality, the metrics tend to agree more.

More information:

- JEG-Hybrid wiki: http://vqegjeg.ugent.be/wiki