# Scenarios for FlexCode

Harald Pobloth

Jan. 10, 2008

# Outline



- FlexCode in Brief
- Scenarios in Brief
- FlexCode Advantage
- Scenarios in some Detail
- Ranking of Scenarios
- Existing Codecs
- Other 6th Framework Projects
- Summary / Conclusions

# FlexCode in Brief

- Exploit
  - flexible signal models, e.g. GMMs
  - high-rate optimized quantization, e.g. SQ or LVQ
  - source-channel decoding
- to obtain coding scheme
  - adjustable to a wide range of rates
  - variable delay
  - variable protection of bit payload on channel
- that serves
  - wide range of services and scenarios

# FlexCode in Brief



- Operates on the continuum of rates
  - intended optimal performance between 10 and 60 kbps
- Computational complexity independent of rate
- No storage of codebooks
  - quantization designed (high-rate approximations) according to
    - available rate
    - source distribution
  - store GMMs to represent signal statistics
    - independent of rate
- Contains advanced perceptual model
- Adapts to feedback from transmission channel
  - Protection strength
  - Rate distribution between channel- and source-coding

FlexCode

# FlexCode in Brief

ERICSSON

- FlexCode organized in work-packages:
  - WP1: Source coding
  - WP2: Channel coding
  - WP3: Real-world scenarios
  - WP4: Integration and demonstrator
  - WP5: Testing
  - WP6: Dissemination and standardization
- WP3 Real-world scenarios provides
  - context and constraints to FlexCode coding
  - list of service scenarios, which benefit from FlexCode

FlexCode

# Scenarios in Brief

- Eight scenarios
- Ranked according to several criteria

1. Mobile Multimedia Blogging Scenario (MMBS)
   - Mobile user records or broadcasts audio-visual content
   - Different rate constraints for record and broadcast case
   - Content consumed life or downloaded from blog-server
   - Content consumed on different devices (TV, PC, mobile phone)

FlexCode

# Scenarios in Brief

ERICSSON

1. Mobile Multimedia Blogging Scenario (MMBS)
2. Multimedia Conference Scenario (MCfS)
   - Users at a variety of locations
   - 3D sound rendering enhances conferencing experience
   - Legacy equipment (mono, stereo, narrowband) support
   - Speech and music content

# Scenarios in Brief

1. Mobile Multimedia Blogging Scenario (MMBS)
2. Multimedia Conference Scenario (MCfS)
3. Mobile Conversation Scenario (MCvS)
   - Much like current circuit switched telephony
   - Conversational
     - Delay
   - Mainly speech content
   - At least one user mobile
     - Battery, computation, bandwidth constraint

FlexCode

# Scenarios in Brief

1. Mobile Multimedia Blogging Scenario (MMBS)
2. Multimedia Conference Scenario (MCfS)
3. Mobile Conversation Scenario (MCvS)
4. Internet Conversation Scenario (ICS)
   - Much like the MCvS
   - Scenario assumes no mobile devices
     - Maximum computational power, high bandwidth
   - Wireless link might be present
     - Necessity to adapt to different error patterns

# Scenarios in Brief



1. Mobile Multimedia Blogging Scenario (MMBS)
2. Multimedia Conference Scenario (MCfS)
3. Mobile Conversation Scenario (MCvS)
4. Internet Conversation Scenario (ICS)
5. Multimedia On-Demand Streaming Scenario (MMSS)
   - User selects content from server
   - Stream is exclusive to user (unicast)
   - Stream is real-time
     - Some delay constraint
   - Both mobile and stationary devices considered

# Scenarios in Brief

1. Mobile Multimedia Blogging Scenario (MMBS)
2. Multimedia Conference Scenario (MCfS)
3. Mobile Conversation Scenario (MCvS)
4. Internet Conversation Scenario (ICS)
5. Multimedia On-Demand Streaming Scenario (MODSS)
6. Multimedia Multicast-Streaming Scenario (MMSS)
   - Stream to several users
   - Real time
   - Both mobile and stationary terminals

# Scenarios in Brief

1. Mobile Multimedia Blogging Scenario (MMBS)
2. Multimedia Conference Scenario (MCfS)
3. Mobile Conversation Scenario (MCvS)
4. Internet Conversation Scenario (ICS)
5. Multimedia On-Demand Streaming Scenario (MODSS)
6. Multimedia Multicast-Streaming Scenario (MMSS)
7. Multimedia Download Scenario (MDS)
   - Not real-time
     - No delay requirements
     - Out of sequence reception possible
   - Unicast
   - Mobile and stationary devices

# Scenarios in Brief



1. Mobile Multimedia Blogging Scenario (MMBS)
2. Multimedia Conference Scenario (MCfS)
3. Mobile Conversation Scenario (MCvS)
4. Internet Conversation Scenario (ICS)
5. Multimedia On-Demand Streaming Scenario (MODSS)
6. Multimedia Multicast-Streaming Scenario (MMSS)
7. Multimedia Download Scenario (MDS)
8. Surveillance Scenario (SuS)
   - Audio surveillance
   - High compression rate while maintaining intelligibility
   - Severe noise might be present
   - Equipment should be cheap
     - Computational constraints

FlexCode

# Scenarios in Brief

ERICSSON

1. Mobile Multimedia Blogging Scenario (MMBS)
2. Multimedia Conference Scenario (MCfS)
3. Mobile Conversation Scenario (MCvS)
4. Internet Conversation Scenario (ICS)
5. Multimedia On-Demand Streaming Scenario (MODSS)
6. Multimedia Multicast-Streaming Scenario (MMSS)
7. Multimedia Download Scenario (MDS)
8. Surveillance Scenario (SuS)

- Others …
  - e.g. gaming, push-to-talk, voice mail

FlexCode

# FlexCode Advantage (General)

ERICSSON

- Device contain more than one services
- Same codec can serve a magnitude of services
- This coder provides savings in
  - licensing costs
  - storage
  - implementation costs

# FlexCode Advantage



1. Mobile Multimedia Blogging (MMBS)
   - Maximum exploitation of bottleneck upload channel
     - Upload channels of different types and at different loads
     - Codec needs to adapt to varying channel characteristics
   - Life vs. upload gives different requirements for codec
     - Delay, rate, error patterns different for life and upload cases
   - Rendering on heterogeneous devices
     - Downscaling at receiving device can save battery / computation
     - Sender can adjust if audience is known (no server upload)

# FlexCode Advantage



2. Multimedia Conference (MCfS)
   - Content variation
     - Varying type and strength of background noise
     - Music or pure speech
     - Single or multiple speakers
   - Varying number of participants
     - Network load increases with increasing # of participants
     - FlexCode can adapt to the varying network
   - Varying number of active speakers
     - Both network load and signal characteristics change with # of active speakers
   - Participants with different terminal / network capabilities
   - Life encoding → network feedback can be used in encoder

# FlexCode Advantage



3. Mobile Conversation (MCvS)
   - Different and varying network capabilities / conditions
   - Life encoding and utilization of feedback
   - Multiparty conversation
     - Can be engaged / disengaged during one session
     - Similar to MCfS
   - Environment noise varies
     - Perceptual model assures high quality speech remains even in noisy conditions

# FlexCode Advantage



4. Internet Conversation (ICS)
   - Same as MCvS
   - More content variation likely (e.g. background music)

# FlexCode Advantage



5. Multimedia On-Demand Streaming (MODSS)
   – Varying content
   – Adaptation to receiver characteristics
     • Uni-cast stream: Can be encoded to the specs of the receiver
     • Error / delay tradeoff for device at hand
   – Different QoS requirements
     • Seemless change from e.g. pre-view to paid service

# FlexCode Advantage



6. Multimedia Multicast Streaming (MMSS)
   - Similar to MODSS
   - Optimized error rate / delay tradeoff
   - Some advantages require embedded stream in MMSS
     - Adaptation to receiver
     - Encoding work load optimized
       - Not necessary to run a multitude of encoders at server

FlexCode

# FlexCode Advantage

ERICSSON

- Multimedia Download (MDS)
  - Adaptive error correction
    - Avoid need to re-transmit
  - Adaptation to receiving device

# FlexCode Advantage



- Surveillance (SuS)
  - Background characteristics
  - Channel conditions
  - Rendering/storing
    - Devices can be selected to temporarily provide higher quality for active rendering / monitoring

# Scenarios in some Detail



- Scenarios described by:
  - Outline
  - User perspective (content, quality)
  - Equipment (user devices, middleware)
  - Networks
  - Requirements
    - Rate
    - Delay
    - Error-robustness
  - Existing codecs
  - Benefit of FlexCode to scenario
  - Standardization relations
  - Commonalities with other scenarios

| Scenario | User perspective | | Equipment | | Network | | |
|---|---|---|---|---|---|---|---|
| | Content | Quality | Sender | Receiver | Sender | Transport | Receiver |
| 3.1 MMBS | Speech, audio, background noise, often simultaneously | High-quality acceptable on home devices. Stereo signals, some mobile devices provide mono only. Bandwidth: 16-32 kHz | Mobile device (phone, PDA including still camera, optionally connected to digital video camera) | Diversity of devices: Mobile, PC, high-quality audio devices (e.g., 5.1 channel setup) | 3G WCDMA uplink radio access, WiFi | Packet core networks, Internet | 3G WCDMA downlink radio access, WiFi or Ethernet |
| 3.2 MCfS | Speech, background noise, multiple speakers, audio | High-quality, Bandwidth: 8-24 kHz | Stationary device, mono input. (See 3.2) | Stationary device with multiple speaker setup, e.g., 5.1. | Ethernet or high-rate WiFi | Packet core networks, Internet | Ethernet or high-rate WiFi |
| 3.3 MCvS | Speech, background noise and audio | High-quality, mono, Bandwidth: 8 kHz or more | Mobile phone | Mobile or fixed phone | 3G uplink radio access, WiFi | PSTN/NGN, core networks, general Internet | 3G downlink radio access, PSTN/NGN for fixed receiver, WiFi |
| 3.4 ICS | Mainly speech, audio should be supported | High-quality, mono, Bandwidth: 8 kHz or more | PC or WiFi phone | PC or WiFi phone | xDSL, optic fibers, Ethernet or WiFi | Internet | xDSL, optic fibers, Ethernet or WiFi |
| 3.5 MODSS | Mixed Content (Speech, Noise, Music), Audio | High-quality, mono but mostly stereo, multi-channel, Bandwidth: 8-24 kHz | Streaming server | Diversity of devices: Mobile (phone, WiFi), dedicated audio / video players, HDTV screens | | Packet core networks, Internet | Low to high-speed downlink radio access, xDSL, optic fibers |

| Scenario | User perspective | | Equipment | | Network | | |
|---|---|---|---|---|---|---|---|
| | Content | Quality | Sender | Receiver | Sender | Transport | Receiver |
| 3.6 MMSS | Mixed Content (Speech, Noise, Music), Audio | High-quality, mono but mostly stereo or multi-channel, Bandwidth: 8-24 kHz | Streaming server | Diversity of devices: Mobile (phone, WiFi), dedicated audio / video players, HDTV screens | | Packet core networks, Internet | Low to high-speed downlink radio access, xDSL, optic fibers |
| 3.7 MDS | Mixed Content (Speech, Noise, Music), Audio | High-quality (FM-radio or DVD quality), mostly stereo or multi-channel, Bandwidth: 16-24 kHz | Content server | Diversity of devices: Mobile (phone, WiFi), dedicated audio / video players | | Core and access networks (FTTH, xDSL, 3G, ...), Internet | Ethernet, WiFi, ad-hoc network, e.g., Bluetooth |
| 3.8 SuS | Speech distorted with background noise | Medium for monitoring, low for storing | Low power, los CPU hardware | Monitoring device (PC, dedicated hardware), storing device | WiFi, Bluetooth | Packet based core network | Ethernet connection |

| Scenario | Requirements | | | Existing Codecs | FlexCode advantage |
|---|---|---|---|---|---|
| | Rate | Delay | Error-rate | | |
| 3.1MMBS | 40-60 kbps | Limited only by device capability ➔ a few hundred ms | Service usable at PLR > 8% | AMR-WB+, e-AAC+ | •Maximum exploitation of upload channel<br>•Rendering on heterogeneous devices<br>•Source / rendering device mismatch |
| 3.2 MCfS | ≈ 24-60 kbps | 200 – 400 ms end to end ➔ ≈ 25 ms algorithmic | Service usable at ≥ 3% PLR | AMR-WB, ITU-G.722.1, ITU-G.722.1.C, G.729.1 | •Content variation<br>•Varying number of participants<br>•Varying number of active speakers<br>•Different network and terminal capabilities to different participants<br>•Life encoding and utilization of feedback<br>•Conference recording with reduced data-rate |
| 3.3 MCvS | ≈ 10-32 kbps | 100 - 300 ms end-to-end ➔ ≈25 – 40 ms algorithmic | ≥ 1% FER, ≥ 8% PLR if transport via Internet | AMR, AMR-WB, EVRC, VMR-WB, EVRC-WB, G.729.1 | •Different network and terminal capabilities / conditions<br>•Exploitation of possible feedback<br>•Multiparty conversation<br>•Adaptation to environment noise |

| Scenario | Requirements | | | Existing Codecs | FlexCode advantage |
|---|---|---|---|---|---|
| | Rate | Delay | Error-rate | Existing Codecs | FlexCode advantage |
| 3.4 ICS | 10-60 kbps | 200 – 400 ms end-to-end ➔ ≈ 25 ms algorithmic delay | Service usable at PLR > 8% | AMR-WB, proprietary codecs, e.g. iLBC or iSAC | •Varying network qualities<br>•Content variation<br>•Exploitation of feedback<br>•Multiparty conversations |
| 3.5 MODSS | BW [kHz]/ rate:<br>4 / 8-16 kbps<br>8 / 12-32 kbps<br>16/14-56 kbps<br>24/16-64 kbps | Limited only by device capability ➔ a few hundred ms | Low PLR of ≈ 1-2% due to re-transmit | AMR, AMR-WB, AAC, HE-AAC v2, AMR-WB+, Windows Media, MPEG Surround | •Adaptation to content characteristics<br>•Adaptation to receiver characteristics<br>•Optimized error rate / delay tradeoff<br>•Different QoS requirements |
| 3.6 MMSS | BW [kHz]/ rate:<br>4 / 8-16 kbps<br>8 / 12-32 kbps<br>16/14-56 kbps<br>24/16-64 kbps | Limited only by device capability ➔ a few hundred ms | Service usable at PLR of ≥ 5%, re-transmit should be avoided | AMR, AMR-WB, AAC, HE-AAC v2, AMR-WB+, Windows Media, BSAC, AAC+, MPEG Surround | •Adaptation to content characteristics<br>•Adaptation to receiver characteristics<br>•Optimized error rate / delay tradeoff<br>•Optimized encoding work load |
| 3.7 MDS | BW [kHz]/ Rate:<br>16/14-56 kbps<br>24/16-64 kbps | Limited only by device capability ➔ a few hundred ms | No errors, re-transmission of lost packets | MP3, AAC, MPEG Surround, Windows Media Technologies | •Adaptive error correction<br>•Adaptation to receiver characteristics |
| 3.8 SuS | 4-10 kbps | ≤ 100 ms to minimize device complexity | PLR up to 10% due to wireless link, re-transmit should be avoided to minimize complexity | AMR-NB, AMR-WB | •Adaptation to background characteristics<br>•Adaptation to channel conditions<br>•Adaptation to rendering / storing device |

# Ranking of Scenarios



- Final ranking according to:
  - Economical relevancy:
    - Operator interest
    - Manufacturer interest
  - End-use interest
  - Degree of novelty:
    - How much can scenario can from FlexCode
  - Ease of implementation
    - How feasible is an implementation within FlexCode project
  - ➢ 5 criteria
  - Scale and ranking points
    - Very low = 0, low = 1, medium = 2, high = 3, very high = 4

# Ranking of Scenarios

- Final ranking table

| | Operator interest | End user interest (general showcase) | Manufacturer interest | Degree of novelty (FlexCode advantage) | Ease of implementation | Ranking points |
|---|---|---|---|---|---|---|
| 3.1 MMBS | Medium | High | High | Medium | Medium | 12 |
| 3.2 MCfS | High | High | Medium | High | Medium | 13 |
| **3.3 MCvS** | **Very High** | **High** | **Very high** | **High** | **Low** | **15** |
| 3.4 ICS | Medium | High | Medium | High | Low | 11 |
| **3.5 MODSS** | **High** | **Very high** | **High** | **High** | **Medium** | **15** |
| 3.6 MMSS | Very High | High | High | Medium | Low | 13 |
| 3.7 MDS | Medium | High | Medium | Low | Medium | 10 |
| 3.8 SuS | Low | Medium | Low | Medium | High | 9 |

# Existing Codecs

| Type | Rate [kbps] | Delay [ms] | Bandwidth [kHz] | Codecs considered |
|---|---|---|---|---|
| Speech, conversational | 6.6-23.85 | 25 | 0.05-7 | AMR-WB |
| | 8-32 | ≈ 48 | 0.05-4 and 0.05-7 | G.729.1 |
| Speech & Audio, conferencing and streaming | 24, 32 | 40 | 0.05-7 | G.722.1 |
| | 24, 32, 48 | 40 | 0.05-14 | G.722.1 Annex C |
| Speech & Audio non-conversational | 6-36 (mono) 7-48 (stereo) | 100-200 | Varying with bit-rate from 6.2 to 19 | AMR-WB+ |
| | 10-44 (mono) 16-52 (stereo) | | Varying with bit-rate from 10 to 17 | 3GPP e-AAC+ |
| Audio non-conversational | Typically between 6 and 64 for mono (12-128 for stereo) | Depending on bit rate | Up to 0.02-20 Depending on bit rate | AAC |
| | | Depending on bit rate | Up to 0.02-20 Depending on bit rate | BSAC |

# Standardization Bodies

- 3GPP
  - Focus on SA4
  - Multimedia Telephony Service for IMS (MTSI)
- ITU
  - Focus on SG 16 WP 3
    - Q23: New speech and audio codec (ITU-T G.MMCC)
    - Q10: Maintainance / extensions of existing codecs
    - Q9: VBR-EV codec development
- MPEG
  - Focus on exploration work on Speech and Audio Coding

# Other 6th FP projects

- Enthrone
  - Streaming and download for mass-market
  - Relies on:
    - MPEG-21
    - Universal multimedia access (UMA)
  - Scalable codecs needed
  - Choice of audio codec still open

# Other 6th FP projects

- ISIS / DANAE
  - Multimedia content search and delivery
    - User interaction
  - Multimedia adaptation
    - Context adaptation (MPEG-21 DIA)
    - DRM
    - Scalable codecs

# Other 6th FP projects



- ARDOR
  - Adaptive rate-distortion optimised sound coder
  - Rate-distortion controlled combination of different coding techniques
  - High-quality audio
  - Overlap in targets and tools with FlexCode
  - FlexCode:
    - More complete approach:
      - Channel coding included
      - Covering wider range of scenarios
      - More scenario oriented
    - Further utilization of high-rate theory in source codec:
      - Eliminate need of storing CB tables
    - Targets lower rates

# Other 6th FP projects



- M-Pipe
  - Cross layer optimization
    - Layer independent descriptor
  - Requires specific network structure
  - FlexCode: Audio source coding more fundamental
  - FlexCode: Channel coding closer to source coding

# Summary / Conclusions



- Eight scenarios described
  - Most scenarios focus on packet switched networks
  - Advantage of FlexCode for scenarios identified
    - Flexible codec gives intrinsic advantage for scenarios
- Flexible codec can serve several scenarios
- Easy adaptation to channel / equipment necessary
  - Target device not known when encoding
  - Content servers serve large number of users
- Two highest ranked scenarios
  - Mobile conversation scenario
  - Multimedia on demand streaming scenario

# Summary / Conclusions

- Overlap with other scenarios identified
  - MCvS large overlap with ICS and MCfS
    - ICS more computational power
    - MCfS focus on multi-party conversation, architecture different
  - MODS some overlap with MMSS and MDS
    - MMSS different architecture (mutlicast vs. unicast)
    - MDS different requirements (unlimited re-transmit, delay)
- Performance of benchmark codecs shown
- Perspective of FlexCode
  - Standardization bodies
  - Other 6th FP projects
- Full document at: http://www.flexcode.rwth-aachen.de/materials.html