

초고선명 디지털 TV 를 위한 차세대 라우드스피커 레이아웃

이영우, 김선민
 삼성전자, DMC 연구소
 {ywbme.lee, sunmin21.kim}@samsung.com

Next-generation loudspeaker layout for Ultra High Definition (UHD) Digital TV

Young Woo Lee, Sunmin Kim
 DMC R&D Center, Samsung Electronics Co. LTD

요 약

본 논문에서는 초고선명 디지털 TV 를 위한 차세대 멀티채널 사운드 시스템의 최적의 라우드스피커 레이아웃을 도출하기 위해 다양한 라우드스피커 배치 환경에서 인지 관점의 오디오 음질 주관평가를 실시하였다. NHK 22.2 채널 시스템, ITU-R BS.775-2 표준의 7.1 채널 시스템과, 실감 음향에 가장 중요한 역할을 하는 Top Layer 라우드스피커에 중점을 두고 다양한 신규 레이아웃 구성들을 비교하였으며, 스튜디오에서 믹싱된 콘텐츠와 B-format 레코딩을 멀티채널로 생성한 콘텐츠를 이용하여 주관 평가를 실시하였다. 주관 평가 결과, Top Layer 에 3 개의 라우드스피커를 가지는 10.2 채널 라우드스피커 레이아웃이 평가에서 사용된 전체적인 오디오 음질의 등급에서 NHK 22.2 채널 시스템과 차이를 인지하기 어렵다는 결과를 도출하였다.

1. 서론

입체 음향 재생을 위해 다양한 라우드스피커 레이아웃이 제안되어 왔다[1-4]. 일반적으로, 더 충실한 입체 음향 재생을 위해서는 더 많은 라우드스피커가 사용되어야 하지만, 라우드스피커 수의 증가는 시스템을 복잡하고 비용의 증가를 가져온다. 최근들어, 모든 라우드스피커가 수평면에만 있는 ITU-R BS.775-2 의 7.1 채널뿐 아니라, 수직 방향의 실감 음향 재생을 위해 University of Southern California (USC) 의 Integrated Media System Center 에서 개발된 10.2 서라운드 사운드 시스템[2-3], Top-Middle-Bottom Layer 로 구성된 NHK 22.2 멀티채널 사운드 시스템[4] 등, Top Layer 에 Top Layer 라우드 스피커를 포함하는 다양한 레이아웃이 제안되고 있다. 본 연구에서는 차세대 멀티채널 사운드 시스템의 최적의 라우드스피커 구성을 찾기 위해 다양한 라우드스피커 배치 환경에 따른 인지 관점에서 오디오 음질을 비교하였다.

2. 콘텐츠 생성 방법

다른 라우드스피커 레이아웃 간의 비교를 위해 고려되어야 할 사항 중 하나는 콘텐츠 생성이다. 가장 이상적인 비교 평가는 실제 음향 시나리오와 이를 다양한 레이아웃으로 재생한 경우의 비교이지만, 이는 원 사운드 필드를 청취 환경으로 가지고 오기 불가능하므로 현실화 될 수 없다. 때문에 레퍼런스 시나리오를 가정하게 되고, 레퍼런스 시나리오를 청취

환경으로 가져오기 위한 방법은 레퍼런스 시나리오를 생성하기 위해 많은 라우드스피커를 사용하고, 다른 라우드스피커 레이아웃으로 레퍼런스 시나리오를 재생하는 것이다[5]. 본 연구에서는 이와 유사한 방법으로 접근하여 매우 좋은 음질의 입체 음향을 생성하는 NHK 22.2 채널 라우드스피커 시스템을 레퍼런스 시나리오로 가정하고, 더 작은 수의 라우드스피커 시스템으로 재생하여 인지 관점에서 오디오 음질을 비교하였다.

22.2 채널 시스템과 다른 라우드스피커 시스템에서의 콘텐츠를 생성하는 방법은 다양한 레이아웃에 가장 좋은 음질을 가지는 콘텐츠를 생성하는 방법과, 22.2 채널 콘텐츠를 먼저 생성 후, 다른 라우드스피커 레이아웃으로 콘텐츠를 다운믹스하는 두가지 방법이 있다. 전자의 경우 각 라우드스피커 레이아웃에 최적의 음질을 생성할 수 있지만, 자명하지는 않고, 후자의 경우 다운믹스하는 매트릭스에 따라 성능의 차이가 있으며 상관성 있는 위상신호들이 한 라우드스피커에 포함될 경우 콤 필터링이 발생할 수 있다. 이런 다양한 현상들의 영향을 최소화하기 위해, 본 연구에서는 주관평가를 위한 콘텐츠를 생성하기 위해 두 가지 방법을 다 사용하였다.

3. 주관 청취 평가

다양한 관점에서 최적의 라우드스피커 배치를 도출하기 위해 두 가지 다른 청취 평가를 실시하였다. 첫번째 청취 평가는 헬리콥터나 비행기처럼 머리 위에서

지나가는 영화 사운드 트랙의 시나리오 경우인, 이동하는 고도의 Point-like 사운드의 오디오 음질을 다양한 라우드스피커 배치 환경에서 비교 평가하였다. 그 다음으로 두번째 청취평가는 Point-like, 와이드, 앰비언스, 잔향 사운드 등 복합적인 사운드 시나리오의 경우에 오디오 음질을 비교 평가하였다.

청취 평가는 삼성전자 DMC R&D 센터의 표준 청취룸에서 수행되었으며, 이 청취룸의 음향 환경은 ITU-R BS.1116[6]의 요구사항을 만족한다. 레퍼런스 시나리오의 재생을 위해 22 개의 라우드스피커와 2 개의 서버우퍼가 NHK 22.2 채널 시스템의 구성에 따라 Top Layer 9 채널, Middle Layer 10 채널, Bottom Layer 3 채널과 2 개의 서버우퍼 채널로 배치되었다. 각 라우드스피커와 중심 청취 장소와의 거리는 원형 배열로 3 미터이다.

3.1. 방향성 음질 평가

Top Layer 의 라우드스피커는 비행기나 불꽃소리처럼 직접적인 사운드 뿐 아니라, 앰비언스나 잔향을 생성에도 중요한 역할을 한다. 일반적으로 영화에서 로켓, 헬리콥터, 제트기, 새 등과 같이 이동하는 사운드는 빠른 이동속도를 가진다. 사운드가 빨리 이동할 때, 사람의 사운드의 위치를 정확하게 따라가는 능력은 고정되어 있는 사운드에 비해 저하된다. 가상 음원이 빠르게 이동한다면 그 음원이 많은 라우드스피커를 거쳐서 이동한다고 해도, 청취자는 일반적으로 사운드의 시작점과 끝점만 인지하게 된다. 또한 사람은 후방의 고도를 정확하게 인지하는 능력이 전방에 비해 저하되는 경향이 있다[7]. 본 연구에서는 위에서 언급한 실질적인 경우에 Top Layer 라우드스피커의 최적의 구성을 찾기 위한 예비 실험으로써 다양한 Top Layer 라우드스피커 구성에 따른 인지 관점에서 방향성 음질을 평가하였다.

3.1.1. 재생 시스템 레이아웃

9 개의 Top Layer 라우드스피커를 가지는 NHK 의 Top Layer 구성이 레퍼런스 시나리오로 가정하였고, 4 채널, 3 채널, 2 채널의 Top Layer 라우드스피커를 가지는 레이아웃이 비교 평가되었다. Top Layer 라우드스피커가 없는 기존 7.1 시스템이 앵커 시나리오로 사용되었다.

3.1.2. 평가 콘텐츠 및 평가 방법

평가 콘텐츠는 영화 오디오 콘텐츠의 일반적인 사운드인 머리위를 지나가는 헬리콥터 사운드를 사용하였다. 가장 정위된 사운드는 라우드스피커의 고도 방향에서 원형으로 이동하게 생성되었고, 또한 방위각 0°, 90°, 180°에서 직선 방향으로 이동하게 생성되었다. 가상 음원은 10° 스프레드 각도를 가지는 Vector base amplitude panning(VBAP) [8] 방법으로 생성되었다. 이동 속도에 따른 인지되는 음질의 영향을 측정하기 위해 초당 90°, 180°의 다른 각속도가 평가되었다. 10 명의 평가자가 참여하였고, 청취 평가는 MUSHRA [9] 평가 방법과 유사한 방법으로 시행하였다. 평가자는 레퍼런스 시나리오와 각 재생 레이아웃간의 방향성 차이를 평가하였으며, 청취룸의 3 가지 위치(중심, 오른쪽 1 미터, 후방 1 미터)에서 평가하였다.

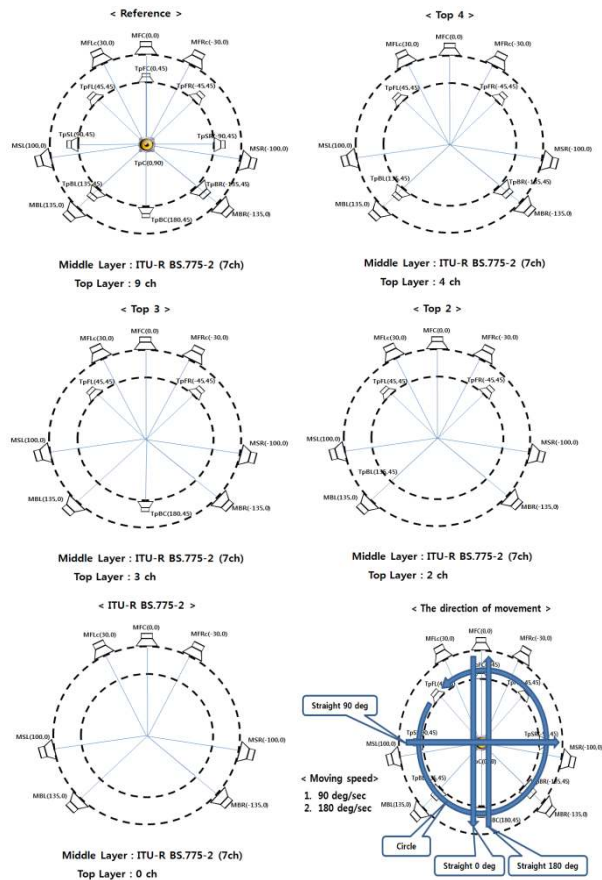


그림 1. 방향성 음질 평가를 위한 라우드스피커 구성 및 사운드 이동 방향과 속도

총실도	등급 범위
인지되지 않음	80-100
인지되나 거슬리지 않음	60-80
약간 거슬림	40-60
거슬림	20-40
많이 거슬림	0-20

표 1. 방향성과 전체적인 음질 평가에 사용된 등급 범위

3.1.3. 평가 결과

각 재생 레이아웃, 청취 위치, 사운드 이동 방향, 이동 속도, 평가자의 독립 변수를 가지는 5 웨이 ANOVA 분석을 실시하였다. 재생 레이아웃 구성이 $F(4,1055)=707.97, p<0.001$ 의 값을 가지며 가장 큰 유의성을 가지는 것으로 분석되어졌고, 각 재생 레이아웃에 따른 신뢰 구간 95%의 평균 스코어는 그림 2와 같다. Top Layer 에 4 개와 3 개의 라우드스피커가 있는 레이아웃은 평균 스코어 80 의 매우 좋은 결과를 얻었고, 2 개의 경우는 약간 나빠졌다. 각 레이아웃간의 유의한 차이성을 보기 위해, Tukey 의 최소 유의차 검증법을 이용한 다중 비교를 시행하였고, Top Layer 에 4 개와 3 개의 라우드스피커 레이아웃 간의 경우를 제외하고는, 모든 재생 레이아웃간에 유의한 차이를 확인하였다. 전반적으로 Top Layer 의 스피커가 증가할수록 오디오 음질이 더 높아지는 것을 확인할 수 있으며, 3 개의 Top Layer 스피커만으로도 매우 좋은 음질을 얻을 수 있다는 결과를 보여주었다.

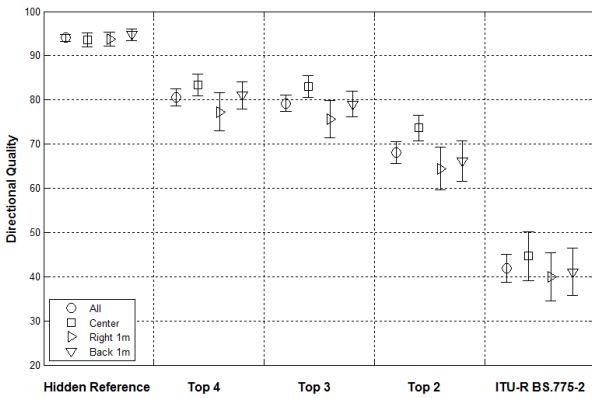


그림 2. 방향성 음질 평가의 다른 수의 Top Layer 라우드스피커 레이아웃에 따른 평균 및 95% 신뢰구간

3.2. 전체적인 음질 평가

이전 청취 평가는 이동하는 사운드의 경우 더 작은 수의 Top Layer의 라우드스피커로 인지할 수 있는 가상 음원의 방향성 음질을 측정하기 위해 설계되었고, 본 청취 평가는 다양한 음원이 포함되는 복잡한 오디오 환경의 재생에서 인지되는 전체적인 오디오 음질에 라우드스피커 레이아웃에 따른 영향을 측정하기 위해 설계되었다.

3.2.1. 재생 시스템 레이아웃

이전 청취 평가에서 좋은 방향성 음질을 얻었던 Top Layer에 라우드스피커가 4개인 삼성 11.2 채널, 3개인 삼성 10.2 채널 레이아웃을, NHK 22.2 채널, USC 10.2 채널, ITU-R BS.775-2의 7.1 채널 레이아웃과 비교 평가하였다. NHK 22.2 채널 레이아웃이 레퍼런스 시나리오로, 7.1 채널 레이아웃이 앵커 시나리오로 사용되었다. NHK 22.2 채널 시스템의 Middle Layer의 Front Left와 Front Right 라우드스피커가 $\pm 60^\circ$ 에 위치하고 있으므로 모든 비교 레이아웃에 같은 프론트 스테레오 이미지를 만들기 위해 삼성 11.2, 삼성 10.2, USC 10.2 레이아웃의 Front Left, Right 라우드스피커의 위치를 동일하게 하였다. 또한 Middle Layer의 Surround Left, Right 라우드스피커의 위치를 NHK 22.2 채널과 USC 10.2 채널의 호환성을 위해 $\pm 100^\circ$ 에서 $\pm 90^\circ$ 로 변경하였다.

3.2.2. 평가 콘텐츠 및 평가 방법

3가지 타입의 평가 콘텐츠가 음향 스튜디오에서 22.2 채널 레이아웃을 위해 제작되었다. 첫번째 타입인 "영화 1" 콘텐츠는 Top Layer에 주로 사운드가 집중되어 있는 총알과 기차가 머리 위로 지나가는 장면으로 구성되어 있는 장면이다. 두번째 타입인 "영화 2"는 기차와 장애물이 충돌하는 장면으로 구성되어 있는 모든 Layer의 라우드스피커로 사운드가 분산되어 있는 콘텐츠이다. 세번째 타입의 콘텐츠는 5.1 채널로 레코딩된 "Eagles - Hotel California"를 믹싱 엔지니어가 4가지 종류의 잔향을 사용해서 22.2 채널로 생성한 음악 클립이다. 생성된 각 22.2 채널 콘텐츠는 다운믹스 매트릭스를 통해 비교 평가될 레이아웃으로 다운믹스 되었다.

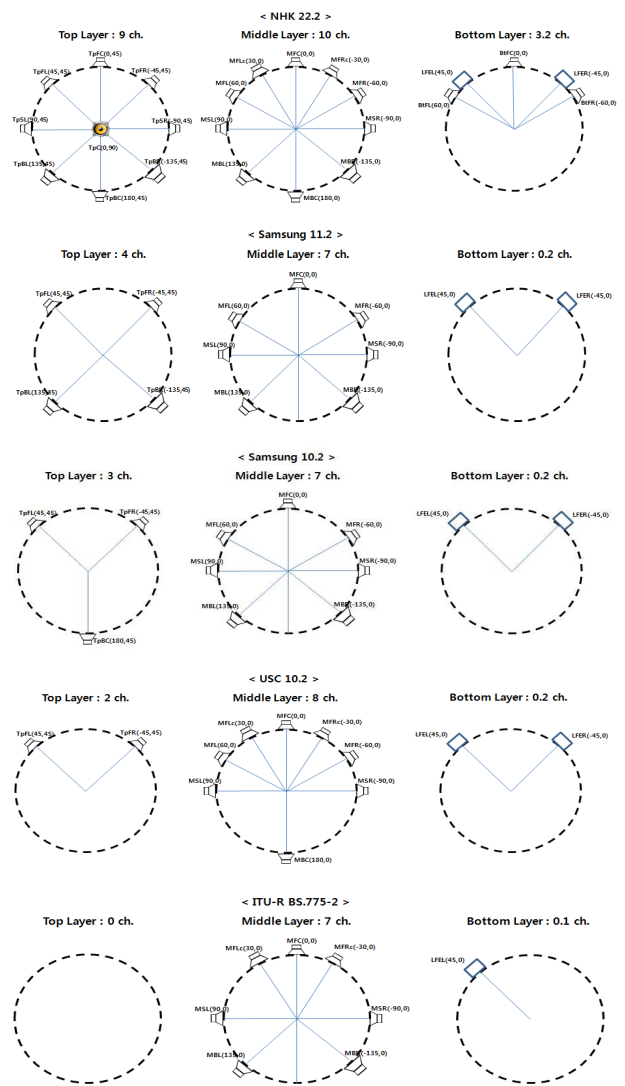


그림 3. 전체적인 음질 평가를 위한 라우드스피커 구성

또한 2 종류의 사운드 클립이 B-format으로 레코딩된 음원을 Directional Audio Coding (DirAC) [5] 기술을 이용하여 각 레이아웃에 맞게 생성되었다. 첫번째 클립은 높은 상공에서 불꽃이 폭발하는 불꽃놀이 장면이며, 두번째 클립은 사람들, 차 이동 소리, 교회 벨소리 등이 레코딩된 도심 광장의 장면이다. 10명의 평가자가 참여하였고, 청취 평가는 MUSHRA 테스트 방법과 유사한 방법으로 시행하였다. 평가자는 레퍼런스 시나리오와 각 재생 레이아웃간의 방향성 차이를 평가하였으며, 청취룸의 3가지 위치(중심, 오른쪽 1미터, 후방 1미터)에서 평가하였다.

3.2.3. 평가 결과

각 재생 레이아웃, 청취 위치, 평가 콘텐츠, 평가자의 독립 변수를 가지는 4웨이 ANOVA 분석을 실시하였다. 방향성 음질 평가에서와 유사하게, 재생 레이아웃 구성이 $F(4,608)=77.91, p<0.01$ 의 값을 가지며 가장 큰 유의성을 가지는 것으로 분석되어졌고, 각 레이아웃에 따른 결과는 그림 4와 같다.

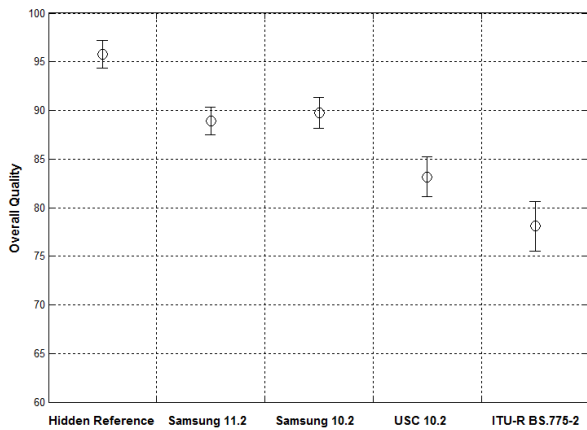


그림 4. 전체적인 음질 평가의 라우드스피커 레이아웃에 따른 인지 차이의 평균 및 95% 신뢰구간

Tukey 검증법에 의해 각 재생 레이아웃간의 유의성 차이를 분석하였고, 삼성 11.2 채널과 삼성 10.2 채널을 제외한 모든 재생 레이아웃간에 유의성 차이를 도출하였다. 평가 콘텐츠의 영향 ($F(4,608)=17.78$, $p<0.01$) 또한 유의성이 있으며, 재생 레이아웃과 평가 콘텐츠 간의 영향도 ($F(16,608)=7.35$, $p<0.01$) 유의한 영향을 가지는 것을 확인하였고, 이는 그림 5 에서 다른 평가 콘텐츠에서 재생 레이아웃에 따른 결과에서 나타난다. 이를 통해 일부 콘텐츠의 경우 다른 요소보다 재생 레이아웃에 따라 큰 차이를 보이는 것을 알수 있고, “영화 1” 콘텐츠의 경우 가장 큰 차이를 생성하였다. 쌍 비교로부터의 결과는 표 2 와 같다. 다양한 라우드스피커 시스템에서 재생되는 평가 콘텐츠들 간의 차이는 비교적 작으며, Top Layer 라우드스피커 개수가 0 개일지라도, 인지되는 음질 차이는 “인지되지만, 거슬리지 않음”으로 평가되었다. “영화 1” 평가 콘텐츠의 경우 머리위에 총알과 기차 같은 직접적인 사운드가 Top Layer 에 집중되도록 구성되어 있으므로 모든 평가자가 쉽게 레이아웃간의 차이를 인지하였다.

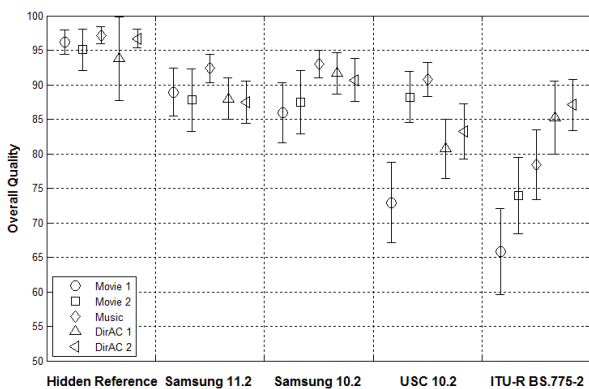


그림 5. 각 평가 콘텐츠에서 라우드스피커 시스템에 따른 전체적인 음질 인지 차이의 평균 및 95% 신뢰구간

Movie 1, System	Hid.Ref.	Top4	Top3	Top2	Worst
Hid.Ref.	○	●			
Top4		○	●		
Top3			○		
Top2				○	●
Worst					○

표 2. “영화 1” 콘텐츠의 다중 비교 결과

4. 결론

본 논문에서는 다양한 Top Layer 라우드스피커 레이아웃 상황에 따른 영향이 평가되었다. 9 개의 Top Layer 라우드스피커를 가지는 NHK 22.2 채널이 레퍼런스 레이아웃이나 실제 환경에 적용하기에는 굉장히 복잡한 시스템이다. 0~4 개의 Top Layer 라우드스피커를 가지는 레이아웃이 NHK 22.2 채널과 비교하여 평가되었고, 평가 콘텐츠는 고도의 가상 음원의 재생, 복잡한 입체 오디오 장면의 재생의 두가지 시나리오로 생성하였다. 결과는 청취자가 NHK 22.2 채널 시스템과 Top Layer 에 3 개의 라우드스피커를 가지는 시스템과의 전체적인 음질 차이를 거의 인지하지 못하였다. 이 결과는 다양한 평가 콘텐츠 및 청취 위치에서도 비교적 일치하며, 앞으로 더욱 다양한 콘텐츠를 이용한 검증이 필요하다.

참고문헌

- [1] ITU-R Rec. BS.775-2, “Multichannel stereophonic sound system with and without accompanying picture,” International Telecommunications Union, Geneva, Switzerland (1992-1994-2006).
- [2] T. Holman, “5.1 Surround sound, Up and Running,” Focal Press, Woburn (2000).
- [3] F. Rumsey, “Spatial Audio,” Focal Press, Oxford (2001).
- [4] K. Hamasaki, et al., “Wide Listening Area with Exceptional Spatial Sound Quality of a 22.2 Multichannel Sound System”, presented at 122nd AES Convention, Vienna, Austria (2007).
- [5] J. Vilkamo, et al., “Directional Audio Coding: Virtual Microphone-Based Synthesis and Subjective Evaluation”, J. Audio Eng. Soc., vol. 57, no. 9, pp.709-724 (2009 Sept.).
- [6] ITU-R BS.1116, “Method for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems,” International Telecommunications Union, Geneva, Switzerland (1997).
- [7] J. Blauert, “Spatial Hearing, The Psychophysics of Human Sound Localisation”, MIT Press (1997).
- [8] V. Pulkki. ”Spatial Sound Generation and Perception by Amplitude Panning Techniques.” Doctoral thesis, Helsinki University of Technology (2001 Aug.).
- [9] ITU-R BS.1534-1, “Method for the Subjective Assessment of Intermediate Quality Level of Coding Systems,” International Telecommunications Union, Geneva, Switzerland (2003).