INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Series P

Supplement 23
(02/98)

SERIES P: TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

ITU-T coded-speech database

Supplement 23 to
ITU-T P-series Recommendations

(Previously CCITT Recommendations)

ITU-T P-SERIES RECOMMENDATIONS

TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series P.10
Subscribers' lines and sets	Series P.30
	P.300
Transmission standards	Series P.40
Objective measuring apparatus	Series P.50
	P.500
Objective electro-acoustical measurements	Series P.60
Measurements related to speech loudness	Series P.70
Methods for objective and subjective assessment of quality	Series P.80
	P.800
Audiovisual quality in multimedia services	Series P.900

For further details, please refer to ITU-T List of Recommendations.

SUPPLEMENT 23 TO ITU-T P-SERIES RECOMMENDATIONS

ITU-T CODED-SPEECH DATABASE

Summary

Supplement 23 to the P series of ITU-T Recommendations is a database of coded and source speech material used in the ITU-T 8 kbit/s codec (Recommendation G.729) characterization tests. The purpose of this database is to provide source, pre-processed and processed speech material, and related subjective test plans and scores, for the development of new and revised ITU Recommendations relating to objective voice quality measures. The coded speech database is delivered on three CD-ROMs.

Source

Supplement 23 to ITU-T P-series Recommendations was prepared by ITU-T Study Group 12 (1997-2000) and was approved under the WTSC Resolution No. 5 procedure on the 27th of February 1998.

FOREWORD

ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

INTELLECTUAL PROPERTY RIGHTS

The ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. The ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, the ITU had received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

Supplement 23 to P-series Recommendations

ITU-T CODED-SPEECH DATABASE

(Geneva, 1998)

1 Background

In the process of standardizing Recommendation P.861 under Question 13/12, it was found that providing a coded-speech database is helpful for developing and/or optimizing objective speech quality measures. After reviewing results of subjective tests carried out in Study Group 12, it was decided to construct the database using the materials from the characterization phase tests for the G.729 speech codec [1][2]. These tests consist of three different experiments. Five organizations agreed to release the speech data: AT&T (USA), CNET (France), CSELT (Italy), Nortel (formerly BNR, Canada), and NTT (Japan). The ITU wishes to thank the organizations for allowing the reproduction of this material.

2 Structure of the database

The database consists of three CD-ROM disks. Table 1 summarizes the contents and structure of the database.

2.1 Document files (Disk #1, directory "\doc")

The test plan [1] and test results [2] are included on Disc #1 as the PDF files "plan.pdf" and "results.pdf" respectively. These files are extremely important since they contain all the information needed to understand the contents of this database.

2.2 Speech files (Disks #1, #2 and #3)

All the speech materials were recorded in 16-bit linear PCM (binary) files with a PC-format (i.e. low-byte first). The filename convention can be found in the test plan [1].

To carry out an objective measurement, it is important to identify the file name of the source/pre-processed speech from the coded-speech file name. For this purpose, Tables 2.1, 2.2 and 2.3 indicate the allocation of talkers and speech samples for each testing condition. The control strings defined in [1] and provided in this database are also helpful.

2.3 Subjective opinion data files (Disc #1, directory "\opinion")

Subjective opinion data is provided in the format specified in [1]. The results of the data analysis can be found in [2].

2.4 Control string files (Disc #1, directory "\lists")

Control strings provide the signal processing information for each of the three experiments. They are recorded in the text format.

2.5 Background noise files (Disc #1, directory "\noise")

Six background noise files used in Experiments #2 and #3 are included in the database. The file format is the same as that of the speech files.

3 README file

The README file, located in the root directory of each of the three disks, is reproduced here:

The speech files contained on this set of disks are the property of the respective test laboratories. Their use has been granted on the condition that they be used to develop new and revised ITU-T Recommendations. Any additional usage must be negotiated with the owner of the data in question.

The files on these disks are organized as follows:

Disk 1 Test plan and results, control string processing lists, additive noise files, individual opinion scores, speech from experiment 1

Disk 2 Speech from experiment 2

Disk 3 Speech from experiment 3

Specific information about the relationship between speech files and processing conditions can be found in the test plan for the relevant test. These test plans are located on Disk #1 in the directory "\doc".

4 References

[1] SQ-46.95R3: *Subjective Test Plan for Characterization of an 8 kbit/s Speech Codec*, Sept. 1995.

[2] TD.66 (WP2/SG15): *Results and Preliminary Analyses of Experiments to Characterize the Subjective Performance of Proposed Rec. G.729*, Nov. 1995.

Table 1 – Directory structure of the database

Disk	Directory			Lab (language)	# of files	Total Mbytes
Disk #1	\doc				2	0.671
	\lists				5	0.040
	\noise				6	5.380
	\opinion	\exp1			3	0.354
		\exp2			3	0.812
		\exp3			4	0.539
	\exp1	\original	\a	CNET (French)	188	48.128
			\d	NTT (Japanese)	188	48.128
			\o	Nortel (English)	188	48.128
		\pre_proc	\a	CNET (French)	188	48.128
			\d	NTT (Japanese)	188	48.128
			\o	Nortel (English)	188	48.128
		\coded	\a	CNET (French)	188	48.128
			\d	NTT (Japanese)	188	48.128
			\o	Nortel (English)	188	48.128
Disk #2	\exp2	\original	\a	CNET (French)	28	7.168
			\d	NTT (Japanese)	28	7.168
			\e	AT&T (English)	28	7.168
		\pre_proc	\a	CNET (French)	144	36.864
			\d	NTT (Japanese)	144	36.864
			\e	AT&T (English)	144	36.864
		\coded	\a	CNET (French)	144	36.864
			\d	NTT (Japanese)	144	36.864
			\e	AT&T (English)	144	36.864
Disk #3	\exp3	\original	\a	CNET (French)	208	53.248
			\c	CSELT (Italian)	208	53.248
			\d	NTT (Japanese)	208	53.248
			\o	Nortel (English)	208	53.248
		\coded	\a	CNET (French)	208	53.248
			\c	CSELT (Italian)	208	53.248
			\d	NTT (Japanese)	208	53.248
			\o	Nortel (English)	208	53.248
		\pre_proc	\a	CNET (French)	208	53.248
			\c	CSELT (Italian)	208	53.248
			\d	NTT (Japanese)	208	53.248
			\o	Nortel (English)	208	53.248

Table 2.1 – Allocation of speech samples for each testing condition (Exp. #1)

Condition	Male 1	Male 2	Female 1	Female 2
1	M01_L01	M02_L2D	F01_L59	F02_L85
2	M01_L02	M02_L2E	F01_L5A	F02_L86
3	M01_L03	M02_L2F	F01_L5B	F02_L87
4	M01_L04	M02_L30	F01_L5C	F02_L88
5	M01_L05	M02_L31	F01_L5D	F02_L89
6	M01_L06	M02_L32	F01_L5E	F02_L8A
7	M01_L07	M02_L33	F01_L5F	F02_L8B
8	M01_L08	M02_L34	F01_L60	F02_L8C
9	M01_L09	M02_L35	F01_L61	F02_L8D
10	M01_L0A	M02_L36	F01_L62	F02_L8E
11	M01_L0B	M02_L37	F01_L63	F02_L8F
12	M01_L0C	M02_L38	F01_L64	F02_L90
13	M01_L0D	M02_L39	F01_L65	F02_L91
14	M01_L0E	M02_L3A	F01_L66	F02_L92
15	M01_L0F	M02_L3B	F01_L67	F02_L93
16	M01_L10	M02_L3C	F01_L68	F02_L94
17	M01_L11	M02_L3D	F01_L69	F02_L95
18	M01_L12	M02_L3E	F01_L6A	F02_L96
19	M01_L13	M02_L3F	F01_L6B	F02_L97
20	M01_L14	M02_L40	F01_L6C	F02_L98
21	M01_L15	M02_L41	F01_L6D	F02_L99
22	M01_L16	M02_L42	F01_L6E	F02_L9A

Condition	Male 1	Male 2	Female 1	Female 2
23	M01_L17	M02_L43	F01_L6F	F02_L9B
24	M01_L18	M02_L44	F01_L70	F02_L9C
25	M01_L19	M02_L45	F01_L71	F02_L9D
26	M01_L1A	M02_L46	F01_L72	F02_L9E
27	M01_L1B	M02_L47	F01_L73	F02_L9F
28	M01_L1C	M02_L48	F01_L74	F02_LA0
29	M01_L1D	M02_L49	F01_L75	F02_LA1
30	M01_L1E	M02_L4A	F01_L76	F02_LA2
31	M01_L1F	M02_L4B	F01_L77	F02_LA3
32	M01_L20	M02_L4C	F01_L78	F02_LA4
33	M01_L21	M02_L4D	F01_L79	F02_LA5
34	M01_L22	M02_L4E	F01_L7A	F02_LA6
35	M01_L23	M02_L4F	F01_L7B	F02_LA7
36	M01_L24	M02_L50	F01_L7C	F02_LA8
37	M01_L25	M02_L51	F01_L7D	F02_LA9
38	M01_L26	M02_L52	F01_L7E	F02_LAA
39	M01_L27	M02_L53	F01_L7F	F02_LAB
40	M01_L28	M02_L54	F01_L80	F02_LAC
41	M01_L29	M02_L55	F01_L81	F02_LAD
42	M01_L2A	M02_L56	F01_L82	F02_LAE
43	M01_L2B	M02_L57	F01_L83	F02_LAF
44	M01_L2C	M02_L58	F01_L84	F02_LB0

Table 2.2 – Allocation of speech samples for each testing condition (Exp. #2)

Condition	Male 1	Male 2	Female 1	Female 2
1	M01_L01	M02_L29	F01_L51	F02_L79
2	M01_L02	M02_L2A	F01_L52	F02_L7A
3	M01_L03	M02_L2B	F01_L53	F02_L7B
4	M01_L04	M02_L2C	F01_L54	F02_L7C
5	M01_L05	M02_L2D	F01_L55	F02_L7D
6	M01_L06	M02_L2E	F01_L56	F02_L7E
7	M01_L07	M02_L2F	F01_L57	F02_L7F
8	M01_L08	M02_L30	F01_L58	F02_L80
9	M01_L09	M02_L31	F01_L59	F02_L81
10	M01_L0A	M02_L32	F01_L5A	F02_L82
11	M01_L0B	M02_L33	F01_L5B	F02_L83
12	M01_L0C	M02_L34	F01_L5C	F02_L84
13	M01_L0D	M02_L35	F01_L5D	F02_L85
14	M01_L0E	M02_L36	F01_L5E	F02_L86
15	M01_L0F	M02_L37	F01_L5F	F02_L87
16	M01_L10	M02_L38	F01_L60	F02_L88
17	M01_L11	M02_L39	F01_L61	F02_L89
18	M01_L12	M02_L3A	F01_L62	F02_L8A
19	M01_L13	M02_L3B	F01_L63	F02_L8B
20	M01_L14	M02_L3C	F01_L64	F02_L8C

Condition	Male 1	Male 2	Female 1	Female 2
21	M01_L15	M02_L3D	F01_L65	F02_L8D
22	M01_L16	M02_L3E	F01_L66	F02_L8E
23	M01_L17	M02_L3F	F01_L67	F02_L8F
24	M01_L18	M02_L40	F01_L68	F02_L90
25	M01_L19	M02_L41	F01_L69	F02_L91
26	M01_L1A	M02_L42	F01_L6A	F02_L92
27	M01_L1B	M02_L43	F01_L6B	F02_L93
28	M01_L1C	M02_L44	F01_L6C	F02_L94
29	M01_L1D	F01_L6D		
30	M01_L1E	F01_L6E		
31	M01_L1F	F01_L6F		
32	M01_L20	F01_L70		
33	M01_L21	F01_L71		
34	M01_L22	F01_L72		
35	M01_L23	F01_L73		
36	M01_L24	F01_L74		
37	M01_L25	F01_L75		
38	M01_L26	F01_L76		
39	M01_L27	F01_L77		
40	M01_L28	F01_L78		

Table 2.3 – Allocation of speech samples for each testing condition (Exp. #3)

Condition	Male 1	Male 2	Female 1	Female 2
1	M01_L01	M02_L33	F01_L65	F02_L97
2	M01_L02	M02_L34	F01_L66	F02_L98
3	M01_L03	M02_L35	F01_L67	F02_L99
4	M01_L04	M02_L36	F01_L68	F02_L9A
5	M01_L05	M02_L37	F01_L69	F02_L9B
6	M01_L06	M02_L38	F01_L6A	F02_L9C
7	M01_L07	M02_L39	F01_L6B	F02_L9D
8	M01_L08	M02_L3A	F01_L6C	F02_L9E
9	M01_L09	M02_L3B	F01_L6D	F02_L9F
10	M01_L0A	M02_L3C	F01_L6E	F02_LA0
11	M01_L0B	M02_L3D	F01_L6F	F02_LA1
12	M01_L0C	M02_L3E	F01_L70	F02_LA2
13	M01_L0D	M02_L3F	F01_L71	F02_LA3
14	M01_L0E	M02_L40	F01_L72	F02_LA4
15	M01_L0F	M02_L41	F01_L73	F02_LA5
16	M01_L10	M02_L42	F01_L74	F02_LA6
17	M01_L11	M02_L43	F01_L75	F02_LA7
18	M01_L12	M02_L44	F01_L76	F02_LA8
19	M01_L13	M02_L45	F01_L77	F02_LA9
20	M01_L14	M02_L46	F01_L78	F02_LAA
21	M01_L15	M02_L47	F01_L79	F02_LAB
22	M01_L16	M02_L48	F01_L7A	F02_LAC
23	M01_L17	M02_L49	F01_L7B	F02_LAD
24	M01_L18	M02_L4A	F01_L7C	F02_LAE
25	M01_L19	M02_L4B	F01_L7D	F02_LAF

Condition	Male 1	Male 2	Female 1	Female 2
26	M01_L1A	M02_L4C	F01_L7E	F02_LB0
27	M01_L1B	M02_L4D	F01_L7F	F02_LB1
28	M01_L1C	M02_L4E	F01_L80	F02_LB2
29	M01_L1D	M02_L4F	F01_L81	F02_LB3
30	M01_L1E	M02_L50	F01_L82	F02_LB4
31	M01_L1F	M02_L51	F01_L83	F02_LB5
32	M01_L20	M02_L52	F01_L84	F02_LB6
33	M01_L21	M02_L53	F01_L85	F02_LB7
34	M01_L22	M02_L54	F01_L86	F02_LB8
35	M01_L23	M02_L55	F01_L87	F02_LB9
36	M01_L24	M02_L56	F01_L88	F02_LBA
37	M01_L25	M02_L57	F01_L89	F02_LBB
38	M01_L26	M02_L58	F01_L8A	F02_LBC
39	M01_L27	M02_L59	F01_L8B	F02_LBD
40	M01_L28	M02_L5A	F01_L8C	F02_LBE
41	M01_L29	M02_L5B	F01_L8D	F02_LBF
42	M01_L2A	M02_L5C	F01_L8E	F02_LC0
43	M01_L2B	M02_L5D	F01_L8F	F02_LC1
44	M01_L2C	M02_L5E	F01_L90	F02_LC2
45	M01_L2D	M02_L5F	F01_L91	F02_LC3
46	M01_L2E	M02_L60	F01_L92	F02_LC4
47	M01_L2F	M02_L61	F01_L93	F02_LC5
48	M01_L30	M02_L62	F01_L94	F02_LC6
49	M01_L31	M02_L63	F01_L95	F02_LC7
50	M01_L32	M02_L64	F01_L96	F02_LC8

ITU-T RECOMMENDATIONS SERIES

Series A	Organization of the work of the ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	**Telephone transmission quality, telephone installations, local line networks**
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	Global information infrastructure
Series Z	Programming languages