Sound-Effect-Vocoder

VSM 201

Bedienungsanleitung
Operating Instructions
Mode d'Emploi

SENNHEISER
3002 WEDEMARK 2
☎ MELLENDORF 05130-8011
TELEX 0924623

Table of Contents

1. Introduction

In the continuing search for new auditory sensations, a large number of electronic devices have been invented during the last twenty years. It is surprising that in all these endeavors for the extension of known sounds the human voice was more or less neglected, although many interesting and unusual sounds can be created, such as speech alienation, articulation of noise and instrumental effect. They may be achieved by a suitable combination of the human voice with two known instruments, the vocoder and a synthesizer or other sound source.

Sennheiser introduces the sound effect vocoder, model VSM 201, designed for multiple applications in recording studios, film and TV productions, for progressive musical groups and entertainers. Many of the possible applications have not even been tried yet, but here are some of the most important examples:

1. Speech or singing from extremely low to very high pitch from simply spoken text, also with unusual vibrato or alienated speech melody.

2. Hoarse whisper made from singing or normal speech.

3. Articulating monophonic or polyphonic instruments.

4. Changing one human voice into another.

5. Multivoice polyphonic singing of a text spoken by one person.

6. Formant shifting (Donald Duck voice) in real time.

7. Voices with disharmonic overtone spectrums.

8. Articulation of natural noise such as talking wind, complaining motor, reminding bells, articulation of a noisy crowd in a sports stadium.

9. Unheard of instrumental effects, e. g. percussive organ.

10. Using the vocoder as a multifilter to select a frequency response

11. Generation of speech-dependent control signals for other equipment, e. g. light organs, synthesizers, etc.

The term vocoder (voice coder) was coined as far back as the thirties by Homer Dudley. The basic function of our vocoder here is characterized by an analysis of the speech which is fed into the unit from a microphone or a tape recorder. In a special filter bank, a spectral analysis is performed and signals are developed which indicate the amount of energy in each channel. At the same time, another analysis section tests if there are voiced or unvoiced sounds at any particular time. The second section of the vocoder consists of a synthesis part which assembles the new synthetic speech from the characteristic filter values mentioned above together with an external replacement signal.

Electronic musical instruments and synthesizers of all makes, especially, the well-known Moog synthesizer, can be used as a source for a replacement signal. The external appearance of the vocoder, VSM 201, is matched in dimensions and design to fit together with the Moog studio equipment.

Within economically reasonable limits, the vocoder contains all possible operating controls and signal outputs in order to allow a multitude of vocoder effects. As a specialty, our unit contains a so-called pause-filling automatic which facilitates the articulation of natural noise and makes talking musical instruments possible.

The most important operating controls are:

1. Level controls for each spectrum channel for correcting the frequency response of the replacement signal and of the input signal and also for adjusting a desirable frequency response for the synthetic speech.

2. Level controls for frequency-depending speech addition, respectively adjustment of the multifilter.

3. Adjustment for the automatic pause-filling through the vocoder part.

4. Adjustment for the automatic pause-filling in bypass position.

5. Defeat switch for the voice-mode analysis.

6. Internal white-noise generator with automatic level adjustment to match the external replacement signal.

All operating controls are located on the front panel to follow a logical sequence. Both signal inputs together with their switches and pilot lights are located on two vertical panels to the left. The controls for the processing are located on the large horizontal panels to the right. The explicit labeling permits the use of the vocoder without having to revert to the manual.

This operating instruction book is designed to assist in achieving certain sound effects which depend on the optimum adjustment of the various controls. It is organized in logical steps which, after only a one-time study, would allow to work successfully with the vocoder, even if you use this instrument only occasionally. As a further help, you will find attached to this manual a number of illustrations showing the vocoder front panel. It is easy to sketch the correct knob positions for certain effects that have been tried out into these illustrations. Using these charts, former desirable effects may be reproduced. Additional front panel charts may be obtained from your Sennheiser agency.

The basic control settings and short experiments described herein shall serve the purpose of explaining the principal possibilities of the vocoder. We recommend to go through the examples shown in this book but then to continue experimenting on your own to create new sound effects.

For best results, it is essential that the operator of the vocoder have two talents: Technical understanding for the electronic circuits, as well as artistic creative capabilities. Sennheiser has found Mr. Heinz Funk, a man who combines these prerequisites in an admirable way. In a cooperation with Mr. Funk and Sennheiser engineers, the Sennheiser VSM 201 was designed. Mr. Heinz Funk who is also the Moog Studio Synthesizer agent takes care of the marketing of the vocoder in East and West Germany, as well as in Berlin.

2. Getting started

2.1 Operational hints

The vocoder is a new electronic instrument and appears to be rather complicated. We recommend to proceed according to the following instructions. In this way, you will become familiar with the various possibilities and you will find it easier to experiment with the vocoder.

It is not necessary for the operator to understand the technical functions of the vocoder completely, although it would facilitate working with the unit. Technical explanations can be found in Chapter 3 of this manual.

First and paramount to remember is:

> A Vocoder has only an Output,
> if there are two Input signals.
>
> Or, put another way:
>
> Speech plus Replacement Signal
> result in an Output.

For the following examples, it is assumed that signal sources such as microphones, synthesizers, tape recorders, and monitoring facilities such as headphones or amplifiers with loudspeakers are available.

Please refer to the fold-out table showing the various front panels. The operating instructions will mention the numbers, as well as the front panel labeling to make the identification as easy as possible.

Panel for SPEECH OR SIGNAL INPUTS

(1) LED Array: logarithmic modulation control for speech and articulation input

(2) LED indicator: automatic gain control

(3) ON/OFF switch for automatic gain control

(4) LED lights up when automatic gain control is switched off.

(5) LED control lights up when the microphone is switched on.

(6) Microphone input switch

(7) Microphone input receptacle for speech or signal. Balanced 200 Ω impedance (1 and 3 - signal, 2 - ground)

(8) Phone jack for line input, speech or articulation signal

(9) Phone jack output for the envelope of the speech or articulation signal

Panel for REPLACEMENT SOUND INPUTS

⑩ LED array: linear modulation control for voiced replacement signal

⑪ LED array: linear modulation control for unvoiced replacement signal

⑫ LED lights up when the detector recognizes the voiced replacement signal

⑬ LED lights up when the detector recognizes an unvoiced replacement signal

⑭ VOICE MODE selector
Middle position: automatic switching of the replacement signal
Right: UNVOICED ONLY
Left: VOICED ONLY

⑮ Selector switch for either voiced replacement signal from the built-in impulse generator or external replacement signal

⑯ Selector switch to choose either unvoiced replacement signal from the internal level-controlled noise generator, or, in position EXTERN, input for unvoiced replacement signal

⑰ Microphone input jack for voiced replacement sound

⑱ PINK FILTER switch in upper position inserts a pink filter into the signal path for the unvoiced replacement signal. In the lower position, the pink filter is off.

⑲ Summing inputs for voiced replacement sound signals

⑳ Summing line inputs for unvoiced replacement sound signals

㉑ Control voltage output for the voice mode detector

㉒ Same as line 21, however, inverted.

VARIABLE CHANNEL EMPHASIS

Panel for Equalizing the Vocoder Synthesis

(23) VARIABLE CHANNEL EMPHASIS, individual level controls for the 20 synthesis channels

(24) The LED lights up when the channel level controls are activated

(25) Vocoder selector switch
Middle position: VOCODER OFF, synthesis out
Right position: VOCODER in normal operating mode, all synthesis channels are operating with preset level
Left position: Level adjustment of all synthesis channels via their individual controls

(26) LED lights up when the vocoder is in its normal mode and the synthesis channels are operating with fixed levels

(27) AC-Power switch: (•) = On ( ) = Off

SPEECH ADDITION/MULTIFILTER

(28) SPEECH ADDITION/MULTIFILTER, individual level controls for mixing of the 20 analysis bands

(29) LED lights up when the speech or articulation signal bypass is activated

(30) BYPASS selector switch
Middle position: BYPASS OFF, both ways
Right position: REPLACEMENT SOUND BYPASS for voiced replacement signals activated
Left position: BYPASS for speech or articulation signal is activated

(31) LED lights up when the replacement sound bypass is activated

(32) BYPASS level controls for both bypass circuits

(33) LED lights up when adding or multifilter is activated through their individual controls

(34) Selector switch for MULTIFILTER OR ADDITION
Middle position: MULTIFILTER OR ADDITION OFF
Right position: Addition via sum control
Left position: Addition or multifilter via individual controls

(35) LED lights up when addition is routed through the sum control

(36) SPEECH ADDITION sum control

SILENCE BRIDGING

Panel with Operating Controls for Pause Filling and Output Jacks

(37) SILENCE BRIDGING, individual controls for the voiced pause-filling signal

(38) LED lights up when the pause filling is routed through the individual controls

(39) BRIDGING selector switch
Middle position: Bridging off
Right position: Pause-filling via sum control
Left position: Pause-filling via individual controls

(40) LED lights up when pause-filling is routed through sum controls

(41) SILENCE BRIDGING: sum control for voiced pause-filling signal

(42) LEDs light up if there is sufficient activity in the individual channels

(43) Output jacks for the channel envelopes of the 20 analysis channels

(44) THRESHOLD level control for the speech/pause detector

(45) LED lights up when the speech threshold detector is recognizing speech

(46) Output volume control

(47) Output jack for a control voltage derived from the speech detector

(48) Unbalanced 600 Ω output

(49) Balanced 600 Ω output

This operating manual is valid for vocoders with serial numbers 13 and higher. Units with lower serial numbers do not have the balanced output, the replacement signal automatic gain control amplifier and the replacement signal bypass. The two first-mentioned circuit improvement can be easily retrofitted, the bypass is much more difficult.

2.2 Basic Control Settings

Before starting to work with the vocoder, it is best to adjust all the controls in the following manner:

(3) AUTOCONTROL switch ON (upper position)

(6) MIC switch OFF (lower position)

(7) MIC INPUT SPEECH OR SIGNAL, connect a 200 Ω microphone

(14) VOICE MODE selector switch to AUTOMATIC position (middle position)

(15) VOICED switch to INTERN (upper position)

(16) UNVOICED switch to INTERN (upper position)

(18) PINK FILTER Off (lower position)

(23) VARIABLE CHANNEL EMPHASIS, individual level controls 1 to 20 adjust to center position

(25) VOCODER selector switch Off (middle position)

(28) SPEECH ADDITION/MULTIFILTER, individual level controls 1 to 20 adjust to middle position

(30) BYPASS selector switch Off (middle position)

(32) BYPASS level control in fully counter-clockwise position

(34) MULTIFILTER OR ADDITION selector switch Off (middle position)

(36) SPEECH ADDITION/MULTIFILTER sum control adjust to center position

(37) SILENCE BRIDGING, individual level controls to counter-clockwise position

(39) BRIDGING selector switch Off (middle position)

(41) SILENCE BRIDGING sum control to counter-clockwise position

(44) THRESHOLD control to counter-clockwise position

(46) VOLUME control somewhat opened (9 o'clock position)

(48) OUTPUT, connect headphone or monitor

(49) Connect mixer, tape recorder, monitor or similar to 600 Ω-output

(27) Power switch On (•), voltage adjustment from 110 V to 220 V is automatic.

Next to the four selector switches (25) (30) (34) (39) and to the microphone switch (6), there are LEDs which are recognizable from a larger distance and signal the mode of operation. The lighting up of these LEDs will not be further mentioned in the forthcoming text.

From now on, whenever we refer to the output, we will only mention output (49). Of course, the signal will be also available at the unbalanced output (48).

2.3 SPEECH OR SIGNAL INPUTS

2.3.1 Signal Path Controls via Speech Bypass

(6) Microphone switch ON (upper position)

(30) BYPASS selector switch to SPEECH (left position)

(32) Turn the BYPASS level control slowly up while talking into the microphone. Leave the control in this adjusted position.

(49) The speech signal, as fed into the input of the vocoder, will be present at the output while totally circumventing the actual vocoder section. This will also check the connections to the monitor.

(6) Microphone switch OFF (lower position)

(8) Feed a high-level signal source, e. g. speech or music from a tape recorder into the input jack for line.

(49) This signal will now appear at the output and permits a check of the connection of the signal source.

2.3.2 Level Control for SPEECH OR SIGNAL INPUTS

The proper function of the vocoder depends very much on the correct adjustment of the various signal levels. For this purpose, the unit is equipped with LED arrays (1) and the AUTOCONTROL-LED (2) for the speech or articulation signal.

Adjust the level of your signal that is connected to the line input (8) in such a way that the AUTOCONTROL-LED (2) will just flicker at the loudest part. Depending on the character of the signal, it is possible for the two red LEDs in the light-emitting diode array (1) to light up. LED (2) will show that the automatic gain control amplifier is operating properly, ensuring the best possible dynamic for the SPEECH OR SIGNAL INPUT.

The automatic gain control for the articulation input can be defeated with switch (3) in order to deliberately overmodulate. If the AUTOCONTROL function is switched off, the warning LED (4) is lit.

After setting the levels for the SPEECH OR SIGNAL INPUTS, switch now

(30) BYPASS selector switch to Off (center position),

(32) BYPASS level control to counter-clockwise position.

2.4 Vocoder Mode

2.4.1 Vocoder Check with Internal Replacement Sound

(25) VOCODER selector switch to NORMAL (right position)

(8) Feed line input with a speech signal from a tape recorder or

(7) use a microphone connected to the microphone input

(6) Microphone switch ON (upper position) when microphone is used

(46) Adjust output volume control to desired volume

(49) The OUTPUT supplies now a vocoded speech signal with the pitch of the built-in pulse generator.

2.4.2 Checking of the VOICE MODE DETECTOR

(14) Switch VOICE MODE selector into UNVOICED ONLY (right position).

Now, the vocoded speech will be derived from the noise signal only the voice sounds "hoarse".

(14) Switch VOICE MODE selector into VOICED ONLY (left position).

Now, the vocoded speech will be derived from the pulse generator only, i. e. unvoiced sounds such as F, SH, S will be generated from the line spectrum of the pulse generator and will sound unnatural.

Please note that all unvoiced sounds must be pronounced in a phonetic way, e. g.

... FFFF ... and not "EFF"
... SSSS ... and not "ESS"

(14) VOICE MODE selector switch to AUTOMATIC (center position).

In this AUTOMATIC position, the VOICE MODE DETECTOR will determine the speech combination. Voiced sounds such as A, E, I, O, U (vowels) will be generated from the internal pulse generator while unvoiced sounds such as F, SH, S will be derived from the internal white noise generator. The momentary position of the voice mode analysis will be signaled by the LEDs and is being indicated by the light-emitting diode (12) for voiced and (13) for unvoiced sounds.

The voice mode analysis has been designed for human speech. Independent of the frequency distribution of the input signal at either (7) or (8), the analysis section switches to either voiced or unvoiced.

The VOICE MODE DETECTOR and LEDs (12) and (13) operate independent of the position of the VOICE MODE selector switch (14).

2.4.3 Vocoder Operation with External VoicedReplacement Signal

(19) LINE INPUT FOR VOICED REPLACEMENT SOUND (both jacks are in parallel). Feed with high-level replacement signal such as a synthesizer, electronic organ, recorded natural sounds, music, etc.

(15) VOICED selector switch to external (EXT.)

(10) Check the LED array for the VOICED replacement signal for modulation in such a way that the lower red LED just starts to flicker at volume peaks. An exact adjustment of the level for the replacement signal is important in order to obtain the optimum dynamic of the vocoder. A built-in AGC amplifier prevents overmodulation. The beginning of compression is indicated by the upper red LED of the array (10).

(30) If you want to listen and to judge the replacement signal, put the BYPASS selector switch into the REPLACEMENT SOUND position (right).
Also

(32) turn up the BYPASS volume control and listen to the replacement sound. After that

(30) switch the BYPASS selector OFF (middle position).

At this point, a decision about the use of the vocoder must be made:

As you know, the vocoder can be used either as a multimodulator or to alienate the speech signal. Both cases will be described in the following. Please go directly to Chapter 2.6 for speech alientation, if the multimodulator application is of no interest at this time.

2.5 Multimodulation

2.5.1 Basic Control Settings for Multimodulation

The signal at the REPLACEMENT SOUND INPUT can be articulated by means of another signal connected to the SPEECH OR SIGNAL INPUT. This can best be explained with the following example:

(19) Connect an organ or synthesizer signal to the LINE INPUT REPLACEMENT SOUND VOICED and play long-sustained chords.

(14) VOICE MODE selector switch to VOICED ONLY (left).

(7) Feed speech into the mike input, or

(8) feed a rythmically accentuated sound signal into the line input for speech or signal. This can be any percussion instrument or even the human voice. For a first experiment, simply knock against the microphone.

(1) Check the modulation with help of the LED array, adjust modulation according to 2.3.2

(10) Check the LED array for the voiced signal and control level according to 2.4.3

(49) The output is now the signal obtained by multimodulation.

The articulation signal connected to the SPEECH OR SIGNAL INPUTS (7) or (8) will be analyzed by the appropriate vocoder circuit and tested with respect to volume and frequency distribution. The measured values (characteristic values of the twenty channels) influence, in the synthesis part of the vocoder, the signal which is fed into the REPLACEMENT SOUND INPUTS (17) (19). In our example, the organ as the replacement signal can only then be heard when an articulation signal is present. During the pauses of the articulation signal, there will be no organ sound audible. Remember the basic rule that you must have a speech input and a replacement signal in order to get an output.

For instance, when the tonal range of the articulation input signal at (7) or (8) is changed (whistle into the microphone), only certain filter channels, according to the pitch of the whistle, will be opened momentarily with the result that only the corresponding partial range of the organ tone will become audible.

On the other hand, if you use a cymbal as an articulation signal, you will notice that its sound spectrum will open other channels simultaneously and that the organ sound will now be formed according to the spectral and time composition of the cymbal.

During the experiments, you will notice that it is also necessary for both signals to coincide as far as their frequency distribution is concerned in order for an output signal to be produced. A high tone at the articulation input (whistle into the microphone) can never result in an audible signal, if the replacement input consists of a low organ bass. Therefore, our basic rule must be extended: Not only is it necessary that there be two signals simultaneously but also that these signals be similar with respect to their spectral composition.

2.5.2 VARIABLE CHANNEL EMPHASIS for Equalizing of the Input Signal

It will now be understandable that a change in the total characteristic of the articulation signal will influence the sound of the replacement signal, if, for instance, the articulation signal at (7) or (8) is rather muffled, i. e. it has not enough high frequencies, the organ sound connected to input (19) can also only be reproduced with few overtones and will therefore sound muffled. This undesirable sound can, however, be altered with help of the VARIABLE CHANNEL EMPHASIS.

(25) VOCODER selector switch to CHANNEL EMPHASIS (left position).

(23) VARIABLE CHANNEL EMPHASIS potentiometers can now be adjusted individually. In our example, channels 1 to 10 should perhaps be individually lowered while channels 11 to 20 could be increased.

Remark: Do not turn all controls (23) to the fully opened position in order to increase the volume, since this will cause overmodulation of the master amplifier. The volume in position VOCODER NORMAL ((25) to the right) and VARIABLE CHANNEL EMPHASIS ((25) to the left) should be approximately the same.

2.6 Speech Alienation

2.6.1 Basic Control Settings for Speech Alienation

(14) VOICE MODE selector switch to AUTOMATIC (middle position)

(15) VOICED selector switch to external (downwards)

(16) UNVOICED selector switch to internal (upwards)

(19) LINE INPUT REPLACEMENT SOUND VOICED must be fed with a suitable high-level input signal.

It is advantageous for the following example to use a broadband signal such as several pulse oscillators of a synthesizer.

(10) LED array for REPLACEMENT SOUND VOICED should be checked for proper modulation, if necessary correct according to 2.4.3

(25) VOCODER selector switch to NORMAL (right position)

(8) LINE INPUT SPEECH OR SIGNAL - feed with speech signal from a tape recorder

(49) The output will now supply the alienated speech. Make further experiments with various sound sources.

It is customary to first record the alienated speech on a tape recorder and then to feed it into input (8) (LINE). Sometimes, it is easier for optimizing the vocoded signal if certain test language is directly fed into the vocoder. For this purpose, use a microphone directly into MIC INPUT (7).

Since the human speech consists of mainly two different sounds:

VOICED and UNVOICED

we will need <u>two</u> replacement signals for the speech synthesis.

In order to give the vocoded speech a certain natural effect, it is advantageous to use a replacement signal for the voiced parts consisting of pulse generators with a dense line spectrum, while for the unvoiced part, a noise signal is best. Also, both signals should be similar with regard to their volume. It is for this reason that the VSM 201 uses an internal noise generator with automatic level adjustment which matches the level of the noise signal automatic to the voiced replacement signal.

For instance, if you reduce the level of the external replacement signal (observe LED array (10)), the level of the internal noise generator will also automatically be reduced, as can be seen by observing LED array (11).

We want to repeat once more that only a broadband replacement signal will result in a speech with high intelligibility. A simple sine wave tone or the chirping of a bird are unsuitable as replacement signals and will not result in intelligible speech.

2.6.2 VARIABLE CHANNEL EMPHASIS (Equalization for the Replacement Signal)

It is frequently the case that the replacement signals do not have enough high-frequency contents and therefore the vocoded speech is somewhat muffled. This can be changed by the VARIABLE CHANNEL EMPHASIS Filters (23) :

(25) VOCODER selector switch to VARIABLE CHANNEL EMPHASIS (left position)

(23) VARIABLE CHANNEL EMPHASIS volume controls - can now be adjusted individually, often as follows:

Channels 1 to 10 individually reduced
Channels 11 to 20 individually increased

Again, it is not possible to turn all controls (23) all the way to the right in order to increase the volume. This will lead to an overmodulation of the master amplifier. The volume should be approximately the same in both positions: VOCODER NORMAL ((25) to the right) and VARIABLE CHANNEL EMPHASIS ((25) to the left).

2.6.3 PINK NOISE FILTER

In order to match the noise signal better to the voiced replacement signal as far as frequency response goes, the vocoder is equipped with a built-in PINK FILTER.

(18) PINK FILTER ON (upper position)

Judge by listening whether or not unvoiced sounds such as F, SH and S will fit better to the voiced sounds.

The VOICE MODE selector switch facilitates a classification of the replacement signals' sound characteristics.

(14) It is easier to judge the sound character of the replacement signals if the VOICE MODE selector switch is alternated between the VOICED ONLY and the UNVOICED ONLY position. After you have made your decision, return the switch to the center position AUTOMATIC.

2.6.4 EXTERNAL UNVOICED REPLACEMENT SOUND

For special alienation effects, it is, of course, also possible to use external replacement signals for the unvoiced parts of the speech.

(16) UNVOICED switch to EXTERNAL (downwards)

(20) LINE INPUT REPLACEMENT SOUND UNVOICED must be fed now with the desired high-level replacement signal

(11) Watch the LED array for UNVOICED REPLACEMENT SOUND. If necessary, correct level according to 2.4.3. The level automatic for the internal noise generator is, of course, not effective now.

(18) PINK FILTER can be switched on, if desired. It will also affect an external unvoiced replacement signal.

2.6.5 Mixing of Replacement Signals

Both line inputs (19) and the MIC input (17) sum voiced replacement signals, while the line inputs (20) sum unvoiced replacement signals. This feature permits the connection of several signal sources without the need for an external mixer. Unusual speech alienation effects can be generated, for instance:

(19) Feed organ chords into one of the LINE INPUTS for the VOICED REPLACEMENT SOUND.

(19) Connect an automatic synthesizer signal to the other LINE INPUT for the VOICED REPLACEMENT SOUND.

2.6.6 MICROPHONE INPUT FOR VOICED REPLACEMENT SOUND

(17) Connect a microphone to MIC INPUT REPLACEMENT SOUND VOICED.

A microphone connected here makes further effects possible. Two examples:

Multimodulation with two acoustically separated sound sources which are picked up by two microphones simultaneously.

(7) Feed the sound of a percussion instrument into MIC INPUT SPEECH OR SIGNAL.

(17) Feed an instrument of the brass, string or woodwind group or a choir into the MIC REPLACEMENT SOUND VOICED.

(14) Switch VOICE MODE selector switch to VOICED ONLY (left position).

Voice exchange between two human voices.

(7) The MIC INPUT for SPEECH OR SIGNAL will receive the articulation, i. e. one narrator talks or sings clearly a text.

(17) The MIC INPUT for the VOICED REPLACEMENT SOUND is fed with the melody, i. e. another voice hums or sings the melody, for instance, just with the sound "A".

(14) VOICE MODE selector switch to AUTOMATIC (center position).

(25) VOCODER selector switch to VARIABLE CHANNEL EMPHASIS (left position).

(23) VARIABLE CHANNEL EMPHASIS volume controls can be adjusted individually for optimizing the sound.

It is possible for a man to talk or sing with the voice of a woman or a child.

2.7 SPEECH ADDITION/MULTIFILTER

The speech addition respectively the multifilter circuits provide additional signal paths from the SPEECH OR SIGNAL INPUTS to the output. In contrast to the SPEECH BYPASS, the multifilter signal is routed through the AUTOCONTROL of the input amplifier and the VOLUME control of the output.

2.7.1 Speech Addition

(34) MULTIFILTER OR ADDITION selector switch to sum control (right position).

(36) SPEECH ADDITION/MULTIFILTER sum control is to be adjusted to the desired volume.

The signal connected to (7) or (8) will now be mixed with the vocoded signal.

If an otherwise very interesting replacement signal does not have enough highs, it is not very suitable for speech alienation. With the VARIABLE CHANNEL EMPHASIS (23), it is mostly possible to correct such a deficiency. A frequency-depending speech addition usually helps.

(34) Switch MULTIFILTER OR ADDITION selector switch to individual controls (left position).

(28) SPEECH ADDITION/MULTIFILTER individual level controls should all be in counter-clockwise position. Then, starting with channel 20, turn up the individual level controls and add some of the high frequency content of the input signal at (7) or (8) to the vocoded output signal. In this example, it may be advantageous to leave channel controls 1 to 12 in their counter-clockwise positions.

2.7.2 MULTIFILTER

Every channel vocoder possesses an analysis filter bank. The VSM 201 allows to use these filters for equalization purposes, independent of the vocoder section.

(25) Set VOCODER selector switch to OFF (center position)

(34) MULTIFILTER OR ADDITION selector switch to individual controls (left position)

(28) SPEECH ADDITION/MULTIFILTER - set controls to marked neutral position.

(7) or (8) - Apply signal to SPEECH OR SIGNAL INPUTS.

(28) SPEECH ADDITION/MULTIFILTER - adjust the individual level controls to obtain desired frequency response.

2.8 Control Signals for External Equipment

A large number of control voltages are generated within the vocoder. Some of these are connected to jacks at the front panel, since they can be utilized to control other associated equipment. All these outputs are short-circuit-proof and the function of the vocoder will not be affected by a short across any of these outputs.

Examples for the control of external equipment:

Switching or pitch changing of synthesizer oscillators

Triggering of circuits at certain sounds

Control of frequency shifters

Light effects, "super light organ"

Important:	All control signals are derived only from SPEECH OR SIGNAL INPUTS (7) or (8).

2.8.1 SPEECH ENVELOPE

The level of the input signal is monitored and a corresponding envelope DC-voltage is available at the SPEECH ENVELOPE (9) jack. The output voltage from 0 V to 7 V has a logarithmic characteristic and corresponds to the perceived loudness of the audio signal as also indicated by the LED array (1).

2.8.2 SPEECH DETECTOR

The speech/pause detector circuit monitors the loudness value and switches a control voltage, depending on an adjustable THRESHOLD (44).

The control voltage is available at jack (47) and supplies:

during pause: + 11 Volt

during speech: + 0.5 Volt

These two voltage levels can be reversed by changing an internal solder connection.

LED (45) lights up when the circuit is in "speech" position.

2.8.3 VOICE MODE DETECTOR

The voice mode detector analyses the input signal with respect to frequency distribution and decides between voiced and unvoiced sounds.

Of course, this applies to non-speech input signals as well; low frequency signals are classified as VOICED, high frequent signals as UNVOICED.

At any one time, one of the following conditions will exist:

Voiced sounds or signals with bass and midrange frequencies:

(12) LED VOICED lights up

(21) Jack VOICED: + 11 Volts

(22) Jack UNVOICED: + 0.5 Volts

Unvoiced sounds, high-pitched signals alone or no signal:

(13) LED UNVOICED lights up

(22) Jack UNVOICED: + 11 Volts

(21) Jack VOICED: + 0.5 Volts

2.8.4 CHANNEL ENVELOPES

The Channel Vocoder VSM 201 has twenty individual channels. The level in each channel is monitored, rectified and the resulting envelope voltages are available at jacks (43) for control purposes. These voltages are proportionate (linear) to the levels and range from 0 to 7 Volts.

LEDs (42) signal the presence of frequencies falling into the range of the respective channel.

(7) Feed a microphone signal into MIC INPUT SPEECH OR SIGNAL

(6) Switch MIC to ON
Speak clearly and sustained the vowel "a" as in "father".

(42) Several channel LEDs will light up indicating the most prominent spectral distribution of this sound (channels 6 ... 11). If you change to a different sound, such as "o", as in "blow", other channel LEDs will light up stronger (probably 1 ... 5).

The areas of strongest energy activities are called formants. Many vowels have several formants, which may differ from person to person. Continue experimenting with other sounds, including consonants such as F, S and SH.

This feature of the vocoder can be used to change the tuning of a synthesizer's oscillator or to control a light organ, depending on the pitch of the input signal, be it speech or music.

2.8.5 Combination of Control Signals

Combining several control voltages to influence external equipment, such as synthesizer VCAs, synthesizer threshold controls and similar circuits, allows to distinguish between even more speech details, e. g. VOICE MODE DETECTOR (22) together with CHANNEL ENVELOPE (42) of channel 19 recognize the consonant "Z", but, with channel 15, the consonant "SH".

2.9 SILENCE BRIDGING with Speech Alienation

The vocoder makes it possible to use natural sounds with human articulation. Some examples are:

Whispering wind

Complaining engin

Reminding church bells

Commentary of a drilling press

Thoughts of a pack of dogs

Articulation of a roaring crowd

Speaking organ, etc.

It is, however, not sufficient to connect the real noise into the replacement input and the speech into the articulation input, as can easily be learned from the following example.

2.9.1 Basic Experiment

We suggest the fabrication of a two-channel tape loop for the following experiment: Record a noise on track I, such as an electric hand drill with different loads, church bells or a similar noise, while on track II, a speech signal should be recorded.

(6) MIC switch OFF (lower position)

(14) VOICE MODE selector switch to VOICED ONLY (left position)

(15) VOICED selector switch to EXTERN

(16) UNVOICED selector switch to INTERN

(19) LINE INPUT REPLACEMENT SOUND is now fed with the noise from the tape, track I.

(8) LINE INPUT SPEECH OR SIGNAL is fed with a speech signal from track II of the tape.

(1) and (10) Check the LED bar indicators and adjust the modulation according to 2.3.2 and 2.4.3

(25) VOCODER selector switch to NORMAL (right position)

(30) BYPASS OFF (center position)

(34) MULTIFILTER OFF (center position)

(39) BRIDGING OFF (center position)

(44) THRESHOLD adjustment in 9 o'clock position

(49) Connect a monitor

Now, you will probably hear a very interesting speech alienation, however, real noise will be difficult to discern and the intended effect will not be quite believable.

The explanation for this is quite simple. The speech consists of very short articulation phases with many pauses between the words and the syllables. The short duration of the speech phases is not sufficient to identify the noise clearly. As we already know, there will only be a signal at the output, if there are two input signals present, namely the speech and the replacement signal. Short noise events from the bells or from the hand drill are simply not sufficient to produce the effect. It is essential for the replacement signal to become audible also during the articulation pauses. This automatic mixing of the replacement signal to the vocoded speech signal is called SILENCE BRIDGING - pause filling.

2.9.2 Adjustment of Pause Filling

In order to produce the desired effect, e. g. a talking drilling machine, it is necessary to match the pause filling signal to the vocoded speech signal. These additional requirements are:

suitable volume

corresponding frequency response

equal phase response

The easiest way to satisfy these conditions is to route the replacement signal during the speech pauses through the filter of the synthesis part of the vocoder. Let us get back to the former example:

(25) VOCODER selector switch to VARIABLE CHANNEL EMPHASIS (left position).

(23) VARIABLE CHANNEL EMPHASIS - adjust the individual volume controls to optimize the vocoded speech.

(39) BRIDGING selector switch to sum control (right position).

(19) Start the tape loop!

(41) SILENCE BRIDGING sum control - slowly increase level until pause-filling effect becomes audible.

The pauses will now be filled with the filtered replacement signal and the intended effect should be readily achievable. There are four methods by which pause filling may further be optimized:

Match volume and equalization of the pause filling signal with the vocoded speech signal (see 2.9.3).

Readjust the THRESHOLD of the SPEECH DETECTOR (see 2.9.4).

Add the VOICE MODE DETECTOR (see 2.9.5).

Vary the volume of the speech signal (see 2.9.6).

The sequence of this optimizing effect can be chosen arbitrarily, however, we will describe this procedure in the above sequence.

2.9.3 Equalization and Volume of Pause-Filling Signal

(39) BRIDGING selector switch to individual controls (left position).

(37) SILENCE BRIDGING individual controls should all be brought into the same position as the sum control (41).

(8) Switch off track II of the tape recorder or simply pull the connector at LINE INPUT (8). Now, you will only hear the pause-filling signal. Note the sound impression and judge whether it is dull or bright compared to the vocoded speech.

(8) Switch on track II, respectively reconnect the plug.

(39) BRIDGING selector switch to OFF (center position). Now you will hear the already optimized speech signal only. As before take note of the sound impression.

(39) BRIDGING selector switch again towards the individual controls (left position).

(37) SILENCE BRIDGING individual controls are now adjusted individually in order to match the pause-filling signal to the speech signal with respect to volume and frequency response.

This entire procedure should be later repeated when all other possibilities of optimizing pause-filling have been tried.

2.9.4 SPEECH DETECTOR When Pause-Filling is Used

As previously mentioned, the speech/pause detector measures the typical volume of the speech input signal and decides, up to a certain adjustable volume, what will be regarded as pause and what is speech.

In the pause mode (LED (45) is not lit), the pause-filling signal will be mixed with a preselected volume (37).

In the speech mode (LED (45) is lit), the pause-filling will be controlled from the characteristic volume of the speech input signal. This control is inversely proportional to the input volume. At large speech levels, no pause-filling will take place and the speech articulation will be unhampered.

With smaller speech levels, a certain pause-filling is present and, therefore, the articulation might suffer. When the speech level decreases even further, the detector will switch to the pause mode with resulting pause-filling without audible articulation.

(44) Vary the THRESHOLD control and take note of its influence.

2.9.5 VOICE MODE DETECTOR When Pause-Filling is Used

For the pause filling, the VOICE MODE SELECTOR switch (14) was this far in position VOICED ONLY.

With certain replacement signals, the unvoiced sounds in the vocoded speech may not be quite satisfactory.

Switching between the chosen replacement signal and a white noise frequently destroys the desired effect. It is, for instance, not conceivable that the harmonics of the bell sound are capable of synthesizing a sharp "s". Therefore, the voice switching of the vocoder must operate differently. In positions BRIDGING-individual controls and BRIDGING-sum control, the hiss or an unvoiced replacement signal will be mixed with the voiced replacement signal, provided the following conditions are met:

VOICE MODE DETECTOR to UNVOICED, (13) lights up

SPEECH DETECTOR to "speech", (45) lights up

BRIDGING selector switch (39) to individual or sum control

(14) VOICE MODE selector switch to AUTOMATIC

(44) THRESHOLD control should be corrected since this will now influence the white noise addition

(18) PINK FILTER can be switched on, judge its effect.

Additional Remark:

The VOICE MODE switch (14) can also be switched to the VOICED ONLY position. In this case, the addition of white noise is always implemented when the SPEECH DETECTOR is in the speech position. These possibilities will be described again in chapter "Addition of unvoiced signals in vocoding mode" (2.11.1).

2.9.6 SPEECH INPUT Level When Pause-Filling is Used

As mentioned in chapter "SPEECH DETECTOR when pause-filling is used", the amount of pause-filling and with that the hiding of the articulation dpends upon the size of the speech input level.

(8) Reduce speech level (track II).

(19) Leave replacement signal at its former level (track I).

Significant pause-filling will now become audible even during the articulation phases and the intelligibility will be reduced in favor of an undisturbed replacement signal.

This effect may also be reversed:

(3) AUTOCONTROL switch OFF, LED (4) lights up as a warning, since this is not one of the usual operating modes.

(8) Increase speech level (track II).

(19) Leave replacement signal (track I) unchanged.

Now, there will be no pause-filling signal during even the low-level articulation phases of the speech, since there will be no pause-filling starting with a certain level. A further increase can therefore not initiate any changes.

Repeat all aforementioned settings for pause-filling and attempt to further optimize the desired effect.

(3) AUTOCONTROL switch - return to its normal position.

2.10 SILENCE BRIDGING for Instrumental Effects

In the previously mentioned examples, we concentrated mainly on the speech intelligibility in conjunction with pause-filling. Choosing pause-filling in instrumental effects, it is important that the sound characteristic of the instrument to be articulated will not be changed in an undesirable way. Therefore, a few deviations from the adjustments are essential.

2.10.1 Pause-Filling Via SILENCE BRIDGING

Firstly, a decision must be made whether or not the instrument signal should be routed through the vocoder synthesis part. Of course, there will be drastic effects with regard to phase shifts, amplitude and frequency characteristics caused by the comb filters of the synthesizer. Adjust the controls as follows for verification:

(19) LINE INPUT REPLACEMENT SOUND VOICED - feed with high-level instrument signal, or

(17) MIC INPUT REPLACEMENT SOUND VOICED - pick up instrument through a microphone.

(3) AUTOCONTROL switch ON.

(14) VOICE MODE selector switch to VOICED ONLY.

(15) VOICED selector switch to EXTERN.

(25) VOCODER selector switch to NORMAL.

(30) BYPASS OFF.

(34) MULTIFILTER OFF.

(39) BRIDGING to sum control (right position).

(41) SILENCE BRIDGING - turn up sum control to approximately 12 o'clock position.

(44) THRESHOLD control - to 9 o'clock position.

Now judge the sound characteristic of the instrument. If the changes in sound can be tolerated, pause-filling can be initiated through the SILENCE BRIDGING circuit, i. e. via the vocoder synthesis.

Further optimizing can be performed with the THRESHOLD control (44), VOICE MODE switch (14) and the VARIABLE CHANNEL EMPHASIS (23) as described before.

2.10.2 Pause-Filling Via BYPASS REPLACEMENT SOUND

Should you wish to observe the sound of the instrument without its being influenced by the vocoder synthesis, proceed as follows:

(39) BRIDGING selector switch to OFF (center position)

(30) BYPASS selector switch to R. SOUND (right position)

(32) BYPASS level control - turned up.

The two sound impressions, via SILENCE BRIDGING and via BYPASS can be compared by means of momentary simultaneous switching of both switches (30) and (39). If there are differences in volume, matching should take place with controls (32) and (41) respectively.

(7) respectively (8) SPEECH OR SIGNAL INPUTS - feed with articulation signal, e. g. with instrument or voice, however, not speech.

As long as this articulation signal is not present, you will hear the instrument. As soon as an articulation signal is fed into the unit, the instrument will become audible in multimodulation mode through the vocoder synthesis. Optimizing of the sound is now possible by switching on VARIABLE CHANNEL EMPHASIS (23) and VOICE MODE selector switch (14). It should be noted that again a switching between VOICED and UNVOICED sounds will take place, as during normal vocoder operation.

2.10.3 Additional Switching Possibilities

If the UNVOICED signal is to be added to the VOICED signal during the UNVOICED passages, proceed as follows:

(39) BRIDGING selector switch to sum cohtrol (right position)

(41) BRIDGING sum control to counter-clockwise position.

(44) THRESHOLD - experiment with the control.

With this setting of the control, the very interesting effect of the speaking organ can be realized. In addition, it is possible to fill the pauses via the BYPASS as well as the SILENCE BRIDGING circuit and further sound variations will result.

In deviation of the regular vocoder function for all pause-filling effects the following applies:

Without an articulation signal being fed into inputs (7) or (8), the replacement signal (R. SOUND) will become audible in the adjusted way. If an articulation signal is applied now, the replacement signal will be reduced in volume depending upon the level of the articulation signal and the replacement signal may be substituted by the modulated signal of the vocoder synthesis.

2.10.4 BYPASS REPLACEMENT SOUND Without Pause-Filling Effect

If the automatic control of the BYPASS R. SOUND is not desired or if the same is too noticeable, a potentiometer at the back of the vocoder permits a reduction of this effect.

2.11 Further Effect Features

2.11.1 Addition of UNVOICED Signals in Vocoding Mode

As we already know, the VOICE MODE DETECTOR analysizes the frequency distribution of the input signal at the SPEECH OR SIGNAL INPUT and switches the two replacement signals according to the recognition. Should it be desirable, for a certain sound effect, to mix the VOICED and UNVOICED during the UNVOICED phase, set the controls as follows:

(7) or (8) SPEECH OR SIGNAL INPUTS - preferably apply a speech signal

(19) LINE INPUT REPLACEMENT SOUND VOICED - feed with a synthesizer signal

(20) LINE INPUT REPLACEMENT SOUND UNVOICED - feed with an additional, e. g. hissing, synthesizer signal

(14) VOICE MODE selector switch to AUTOMATIC (center position)

(15) VOICED SELECTOR switch to EXTERN (lower position)

(16) UNVOICED switch to EXTERN (lower position)

(25) VOCODER selector switch to NORMAL (right position)

(39) BRIDGING selector switch to sum control (right position)

(41) SILENCE BRIDGING sum control to counter-clockwise position

(44) THRESHOLD control - try for best result between counter-clockwise position and 9 o'clock

(49) OUTPUT supplies the above described mixing of the two replacement signals under the following conditions:

(13) LED UNVOICED is lit

(45) LED THRESHOLD signalizes speech

2.11.2 Using an Additional Frequency Shifter

In conjunction with an external frequency shifter, a number of other interesting alienation effects may be achieved. The shifter can be connected to the signal path of the SPEECH OR SIGNAL INPUTS or to the REPLACEMENT SOUND INPUTS with quite a difference in results.

Voice Change Through Shifting of Formants

Adjust shifter to 0 shift.

Connect a speech signal from the tape recorder to the shifter input.

(8) LINE INPUT SPEECH OR SIGNAL - connect to shifter output

(19) LINE INPUT REPLACEMENT SOUND VOICED - connect to a suitable synthesizer output

(14) VOICE MODE selector switch to AUTOMATIC

(15) VOICED switch to EXTERN

(16) UNVOICED switch to INTERN

(25) VOCODER selector switch to NORMAL or also VARIABLE CHANNEL EMPHASIS

(34) MULTIFILTER OFF

(39) BRIDGING OFF

(49) At first the OUTPUT supplies the alienated speech with correct articulation. Now, slowly engaging the shifter will change the character of the voice with increasing shift.

Replacement Signal Modification by Means of a Frequency Shifter

Usually, musical replacement signals are characterized by a harmonic spectrum. It is possible to use a frequency shifter to generate a nonharmonic spectrum which is suitable as replacement signal for the vocoder and will result into a quite unusual effect.

3. Technical Section

3.1 History

In part 2 you have become acquainted with the operation of the vocoder, as well as with the more important vocoder terms and interesting effects. In an effort to even better understand the many applications, it is advantageous that the actual function of the vocoder be understood intimately.

Originally, vocoders were designed for speech communication purposes. The process begins by analyzing the speech and sorting it according to the number of components or characteristic speech values. These characteristic values are then transmitted through a telephone line or a wireless transmission chain to the synthesizer part of the vocoder in which the characteristic values together with a suitable replacement signal are used to generate a synthetic speech with excellent intelligibility, see illustration 3. The advantage of this complicated procedure must be seen in a saving of required bandwidth, justifying the expense for the complicated circuits at both ends of the transmission path in certain instances.

It has been known for some time that various replacement signals may be used to obtain very interesting vocoder effects. However, only with the availability of electronic musical instruments, such as synthesizers and electronic organs, access to a wide variety of replacement signals has become a reality. At this stage, the time had come for the development of special sound effect vocoders, such as Sennheiser model VSM 201.

3.2 Human Voice

For a better understanding of the vocoder principle, it is helpful to take a closer look at human speech generation. The various human sounds may be grouped into two major categories: voiced sounds such as A, E, I, O, U and unvoiced or hissings sounds such as F, SH, S. The unvoiced sounds should be pronounced here in a phonetic way, e. g.

... FFFF ... and not "EFF"
... SSSS ... and not "ESS".

While forming a voiced sound, the vocal cords interrupt the stream of air periodically and generate triangularly-shaped pressure pulses. The number of interruptions per second, respectively the sequence of the pulses, determine the musical pitch of the voiced sound. This is the so-called basic melody or melody frequency. With low-bass voices, this frequency is located at around 80 Hz and may reach up to over 1000 Hz with a female soprano. Normal melody frequencies for males and females encompass a range from about 90 to 400 Hz.

Aside from the fundamental frequency, the pressure pulses from the vocal cords contain many harmonics of up to over 4 kHz which can graphically be shown as a line spectrum in an amplitude versus frequency diagram.

The cavities in mouth, throat and nose areas can be regarded as filters which select certain parts of the vocal spectrum in order to form the typical characteristic harmonics of the individual voiced sounds, see illustration 4 and 5. The frequency response can be easily depicted by an imaginary envelope curve.

Significant frequencies with concentration of energy are called formants. They are very similar for both the female and the male voice. However, on account of the higher female fundamental voice, the spectrograms differ noticeably by larger distances of the harmonics.

The vocal cords are not activated for unvoiced sounds and the air stream can pass unhindered. Hissing sounds are generated by air turbulence stiking mouth, tongue, teeth and lips. They can be thought of as specially filtered white noise with typical envelope curves for each individual unvoiced sound. Here, no characteristic differences between male and female voices exist. Explosive sounds such as "P" will not be discussed at this point.

3.3 Principle of Vocoders

Almost all vocoders have an analysis section, as shown in the block diagram of illustration 6. There are usually three major parts: spectrum analysis, voice mode analysis and melody analysis.

By suitable means, the spectrum analysis determines the real time envelope of speech sounds, according to illustrations 4 and 5. A large number of characteristic spectrum values (channel envelopes) are supplied which can be regarded as coding for the position of mouth, nose and throat filters.

The VOICE MODE DETECTOR determines whether or not vocal cords are oscillating or if these permit the air stream to pass freely, i. e. whether the sound is voiced or unvoiced. The output of the voice mode analysis supplies a Yes/No-signal (VOICED/UNVOICED) indicating the actual throat function at any time.

The melody analysis determines the fundamental frequency and supplies the characteristic melody value, the pitch.

The synthesis is a reversal of the analysis procedure. There are four main sections: spectral synthesis, voice mode switcher, impulse generator and hiss generator.

In the case of a communications vocoder, the basic frequency of the impulse generator is controlled by the characteristic melody value. Together with the hiss generator, an imitation of the human larynx is provided. In contrast, it is not necessary for a sound effect vocoder, which is used for purposes of speech alienation, to analyse the characteristic melody value (pitch), since a voiced replacement signal will be used in lieu of the aforementioned impulse generator. Usually, this signal is taken from a synthesizer or an electronic organ which will now determine the pitch of the synthetic voice.

The voice mode detector will control the voice mode switch which alternates between the voiced replacement signal source and the unvoiced replacement signal source, usually a hiss generator.

The spectrum synthesis is derived from the replacement sound and the spectral envelopes. Thus, the spectrum synthesis is an imitation of the human mouth, nose and throat filters.

3.4 Basics of a Channel Vocoder

A number of different circuits are known to perform the various functions of the building blocks of a vocoder and the development for new circuits is still in progress. The various vocoder types are classified according to their method of spectrum analysis. The vocoder introduced by Dudley belongs to the group of channel or filterbank vocoders. Sennheiser vocoder, model VSM 201, equally belongs to this group.

Signal processing of a channel vocoder is shown in illustration 7. An AGC amplifier supplies the microphone signal to a set of channel filters, the so-called filterbank. Each one of these band pass filters analyzes a small portion of the speech spectrum. The number of filters in actually built vocoders may vary between ten (as in the Dudley vocoder) and twenty-four and more.

The larger the number of filters, the better the accuracy of analysis and quality of reproduction. On the other hand, the circuits become more elaborate and expensive and, in the case of communications vocoders, the required bandwidth will be increased.

Sennheiser's sound effect vocoder contains twenty channels, assuring excellent reproduction quality and reasonable cost. Illustration 8 depicts the frequency response of the individual channel filters. These are followed by two-way rectifiers and a low-pass filter for each channel. The lower frequency of these channel low-pass filters is in the range of 20 Hz to 50 Hz. This low-limit frequency is tolerable since the human speech cannot change faster in view of the relatively slow muscle movements of mouth and throat region. The output voltages of the low-pass filters are the characteristic channel values of the coded speech signal.

It may be helpful to look at some oscilloscope pictures showing the signals at various junctions between the building blocks of illustration 7. The amplified microphone signal, prior to entering the filterbank, is shown in illustration 9.1. What we see here is the beginning of the German way of prononciating the word "bass" similar to the English word "but". The "a (English u)" contains a domineering formant in the range of 500 to 1000 Hz, as per illustration 4. The analysis filter of channel 6, with a middle frequency of approximately 700 Hz, will encompass this range and supply a high output voltage during the sound "a (English u)", as depicted in illustration 9.2. The amplitude modulation caused by the fundamental speech frequency can be easily recognized, approximately 2.5 divisions on the scope screen. Negative half waves are removed by a subsequent rectifier (see illustration 9.3). Then follows the channel low-pass filter which removes the AC contents of the signal and supplies the characteristic channel value, as shown in illustration 9.4. The low-pass filter suppresses the speech channel of 700 Hz to a large extent and will somewhat affect the modulation with the fundamental speech frequency. However, the envelope of the speech channel will remain as desired. Please note the increase in voltage at the beginning of the sound "a (english u)". All other characteristic spectrum values are functioning similarly and will therefore not be further discussed.

The synthesis is a reversal of the analysis procedure. Here, the voiced signal (Illustration 9.2) consists of needle pulses with a repeat frequency corresponding to the fundamental speech frequency, again, approximately 2.5 divisions. The following circuit modulates the amplitude of these pulses, according to the characteristic channel values, illustration 9.6. The modulated impulses contain a broad harmonic spectrum from which each channel filter of the synthesis part will derive a synthetic speech channel (9.7) which will be very similar to the output signal of the analysis filter. Please compare illustrations 9.7 and 9.2. In the following mixer, partial bands from all channels will be added to result in the complete synthetic speech, illustration 9.8. The oscillogram looks quite different from the original speech signal which is caused by phase shifts and certain inaccuracies of the vocoder principle (compare illustrations 9.8 and 9.1). The synthetic speech is still easily understood, since all important formants, essential for intelligibility, are accurately retained.

3.5 Voice Mode Analysis

There are various methods for the voice mode detection. Most modes differ with regard to their spectral distribution. Frequency-depending filtering can be used. Voiced sounds always have a higher level in the frequency range of up to 1.5 kHz, as compared to the range above 4 kHz. The exact opposite is true of unvoiced sounds.

Of the various possibilities for the voice mode detection, one is especially suitable for the vocoder, VSM 201: Our detector circuit compares the characteristic value of a high-frequency channel above 5 kHz with that of a particular low-frequency channel below 1 kHz. The circuit functions very reliably, especially, if there is a broad-band speech signal, as in this case.

3.6 Melody Analysis

In a sound effect vocoder, the pitch of the synthesized voice is usually determined by the replacement signal. Therefore, a melody analysis is not necessary and is mentioned only at this point for the sake of completeness.

The melody analysis determines the actual fundamental frequency out of a very complex speech signal. Again, there are various possibilities to do this. The best-known procedure is based upon direct filtering of the fundamental wave by means of a low pass. On account of the different gliding changes between individual sounds, all known procedures do occasionally cause wrong detection of the fundamental frequency which leads to quality deterioration of the synthetic speech. Therefore, additional logic circuits are required which recognize wrong test results and make the necessary corrections.

3.7 Explanation of Block Diagram

The extensive block diagram of the vocoder, VSM 201, is depicted in illustration 10. It is rather complicated and probably suitable only for those readers who are familiar with the study of block diagrams and who, in addition, are very interested in learning about the technology of the vocoders in extenso. Therefore, the following explanations have been limited to the essentials. In addition, some of the LEDs, switches and potentiometers are not illustrated and labeling, such as B 100 etc., refer to subassemblies of the unit.

Speech or articulation signal are fed to a microphone or line input and subsequently routed through an AGC amplifier which prevents overmodulation of the following vocoder channels. The twenty spectrum channels are arranged in the already known manner: channel filter, rectifier, low pass, modulator, channel filter and mixer. Channel control potentiometers PV 1 to PV 20 are situated behind the synthesizer band filters and thus afford a frequency response equalization of the vocoded signal. The vocoder selector switch SV permit disabling of the vocoder synthesis for operation of the unit as multifilter. The channel envelopes are available at the output jacks KK1 to KK 20. Each one is associated with an LED which starts to light up at approximately 20 dB full modulation of the channel. A final amplifier with volume control PG constitutes the output of the vocoder.

The outputs of the analysis filter supply a second mixer, via multifilter controls PM 1 to PM 20. The amplifier selector switch SM is located before the volume control. The speech signal bypass connects the speech or articulation input with the output, bypassing the volume control. The signal path contains the level control, PB, and the bypass selector switch, SB. The replacement signal bypass is the connection between voiced replacement signal input and volume control. In addition to level control, PE, and bypass selector switch, SE, a modulator for pause-filling is provided. More about pause-filling later.

The voice mode analysis is performed by comparing the characteristic channel value of channel 19 with that of a special channel 0, containing a low-pass instead of a band-pass filter. A comparator is used and its output represents the voice mode. The output and the inverted output of the comparator are accessible at the two voice mode detector jacks, each equipped with a signaling LED. Normally, the voice mode signal controls the automatic voice mode switch which connects the voiced replacement signal to the channel modulators during the voiced phase, and of the articulation input signal, respectively the unvoiced replacement signal during the unvoiced phase of the articulation input signal. The voice mode selector switch, SS, allows to defeat the automatic and to keep either of the replacement signal sources connected at all times.

Mostly, the voiced replacement signal is supplied from an external source. An AGC amplifier prevents overmodulation of the synthesis section. The internal pulse generator merely serves for testing purposes, while the internal hiss generator is very suitable for all regular speech alienation applications. The hiss level is automatically matched to the voiced replacement signal by means of a rectifier and modulator. Frequency distribution of the hiss can be altered with a switchable pink noise filter. For certain effects, an external unvoiced replacement signal may be used. An AGC amplifier is built in for this purpose to supply a proper signal level to the voice mode switcher.

A rectifier without prior filtering produces the speech envelope respectively the characteristic level value. This is connected to a jack, to a logarithmic amplifier for pause-filling and to a comparator, the speech/pause detector. Its output is available direct and inverted. The threshold for speech/pause detector is adjustable (PS). An LED signals speech stage of the detector.

The logarithmic amplifier is followed by an analog inverter which supplies the characteristic value for pause-filling (silence bridging). Its voltage is high during pauses and 0 at strong articulation signals. Via selector switch SA the characteristic pause-filling value is connected to the modulator in the signal path of the replacement signal bypass and, via controls PA 1 to PA 10, to the low-pass filters of the analysis channels.

The speech/pause detector influences the analog inverter in such a way that the characteristic value for pause-filling increases to the maximum voltage during pauses. Associated with the pause-filling selector switch is a circuit with switch SZ which changes the operation of the voice mode switcher. With pause-filling ON, the hard switching between voiced and unvoiced replacement signals is avoided and, instead, the voiced replacement signal stays on while the unvoiced replacement signal is added according to this logic: Voice mode selector switch to AUTOMATIC $\wedge$ voice mode detector: unvoiced $\wedge$ speech/pause detector: speech.

Probably, the explanations in this third part of the manual are difficult for many users. For a successful use of the instrument, it is not really necessary to understand the function of the vocoder in its entirety. However, after mastering the operation as described in part 2, a more profound technical knowledge of the vocoder may be helpful and lead to even more creativity.

4. Technical Data

Vocoder characteristics

Number of filter channels	20
Spectral range	100 Hz ... 8000 Hz

Powering

Automatic selection of line voltage	110 V ± 15% and 220 V ± 15%
Power consumption	appx. 18 Watts

SPEECH OR SIGNAL INPUTS

a. Line input	+ 6 dBm, 100 kΩ, unbalanced
b. Microphone input (may be switched off)	for low-impedance, balanced dynamic microphones (2.5 mV, 5 kΩ)

REPLACEMENT SOUND INPUTS

For voiced replacement signals: two line inputs	0 dBm, unbalanced, 100 kΩ
One microphone input	2.5 mV, unbalanced, 5 kΩ
Mix these three inputs internally.	
For unvoiced replacement signals: two line inputs	0 dBm, unbalanced, 100 kΩ
Mix these two inputs internally.	

Modulation Controls

Input signals	3 LED bar indicators
Characteristic channel values (envelopes)	20 LEDs

Output Signals

Balanced, line (Tuchel)	+ 6 dBm
Unbalanced, line (phone jack)	+ 6 dBm

Control Potentiometers

20 x VARIABLE CHANNEL EMPHASIS, individually for each filter channel

20 x SPEECH ADDITION/MULTIFILTER, individually for each filter channel

1 x Speech addition, simultaneously for all channels

10 x SILENCE BRIDGING, one control for two adjacent channels

1 x SILENCE BRIDGING, simultaneously for all channels

1 x BYPASS for speech, disregarding the volume control, or BYPASS for voiced replacement signal through volume control

1 x Output VOLUME

1 x Speech/pause detector THRESHOLD

1 x Pause-filling characteristic of the BYPASS for voiced replacement signal (control accessible at back panel)

Outputs (1/4" phone jacks)

1 x SPEECH ENVELOPE

2 x VOICE MODE DETECTOR, voiced and unvoiced

20 x CHANNEL ENVELOPES

1 x Unbalanced output at + 6 dBm, 300 Ω minimum load

1 x Balanced output + 6 dBm (Tuchel connector), 300 Ω minimum load

Control switches

Vocoder	OFF, Normal, Formant Control
BYPASS	OFF, Speech Signal, Replacement Signal
SPEECH ADDITION/MULTIFILTER	OFF, Normal, Individual Channels
SILENCE BRIDGING	OFF, Normal, Individual Channels
Voice mode selector	Automatic, Voiced Only, Unvoiced Only
Voiced replacement signal	External, Internal
Unvoiced replacement signal	External, Internal
PINK FILTER	ON/OFF

Dimensions 470 x 260 x 220 mm