REAL TIME SOUND TO NOTE CONVERTER

- AudioToMidi -
(Freeware ver.1.01)

1.0 FEATURES

The present software allows the conversion of a standard audio signal to MIDI signal with an insignificant delay. The resulting MIDI signal can be given to a standard MIDI device, PC speaker, built-in sequencer. The audio signal spectrum is displayed in a special window in real time. The present software also allows special options:

-Input audio signal volume normalization;
-Input audio signal correction by built-in graphic equalizer;
-Consideration of possible audio signal frequency deviation from the standard note frequencies, e.g., because of the difference between the guitar sound and the tuning fork;
-Note sensor selectivity meaning the sensor sensibility of a particular note to the adjacent note signals. The graphic representation of the selectivity is provided. The parameter of this option impacts the conversion delay time.
-Note volume filter allowing to ignore low notes and noise;
-Output note filter, allowing to ignore the notes of pre-set loftiness when MIDI signal generation. The note interval and/or presumed key can also be set.
-Note duration filter allowing to ignore accidental short notes when writing to the built-in sequencer.
-Graphic simulation of the first four harmonics of recognized instrument or voice;
-MIDI signal transposition. Resulting note loftiness shift at the integer number of semi-tones.
-MIDI instrument selection when signal generation;
-Resulting note volume selection.
-MIDI channel selection when MIDI signal generation.

-Monophonic mode which allows the lowest note selection from a number of simultaneous notes, thus providing the separation of the first instrument/voice harmonic, and errorless conversion of monophonic melodies.

The general setting is automatically saved. Separate saving, opening and resetting of the equalizer, harmonic model and filter settings are provided.

The conversion result is presented in real time by highlighted piano keys. The piano keyboard window can also be used to generate the sound corresponding to the pressed key.

The signal spectrum representation allows the software usage for the comfortable guitar tuning. The visual peaks must be symmetrical at the middle position of the "Tune" control slider.

The built-in sequencer allows opening and playing MIDI (*.mid) and RIFF MIDI (*.rmi) files. New tracks can be also added to the open files. A new record in the sequencer is made by addition of a new track. Thus, MIDI record can be created from several tracks. The record can be saved in MIDI or RIFF MIDI files.

The software provides the selection of the input audio device and the output MIDI device. Due to the selection option of the output MIDI device, AudioToMidi is able to operate with an external software sequencer. The driver Sonic Foundry Virtual MIDI Router (VMR) is recommended for this purpose. The distribution conditions of this file with the description in English are found in the enclosed file Sonic Foundry MIDI Router.wri (the driver package is not included). MIDI signal, both real time and written to the built-in sequencer, can be applied to the external software sequencer. In the latter case all tracks are combined into one track.

The audio *.wav, *.mp3, *.au files and the like can be converted by the playback by the appropriate software giving the sound to the audio device selected at the AudioToMidi input (usually Wave Mapper).

2.0 MINIMUM SYSTEM REQUIREMENTS

Processor:	P75.
OS:	Windows 95 or Windows NT.
Memory:	Approximately 1MB free ROM.
Hard disk:	Approximately 1MB free space.
Devices:	Any sound card, which is not worse than SB16 in possibility. Drivers and application programs of the sound card must be installed to OS.
Note:	No sound is given to PC speaker under Windows NT, the appropriate option is disabled.

3.0 INSTALLATION

This software product is installed by simple copying of files to any directory preserving the archive catalog tree structure. No additional libraries or drivers are required. The file AudioToMidi.exe runs the software.

4.0 OPERATION PRINCIPLE

A continuous sample flow representing a digitized sound from the "Wave In" list device is given to the program input. This signal is given to the massive of sensors; each tuned at a particular frequency. This frequency is equal the frequency of the note associated with the sensor, at some possible correction. The value representing the sound intensity within the domain of the sensor

own frequency is generated at the each sensor output. These values are graphically presented in the "Spectrum" window.

The note frequency is commonly calculated. Note A of the first octave has the frequency of 440 Hz. When the note is raised or lowered at 1/2 tone, the frequency is multiplied at or divided by the value, equal to the 12-power root of two. Hence, if the note is raised or lowered at 12 semi-tones, i.e., at one octave, the frequency is multiplied at or divided by 2. A of small octave matches 220 Hz, A of large octave - 110 Hz, A of the second octave - 880 Hz, A of the third octave - 1760 Hz, etc. The correction of the sensor frequency is a function of the position of "Tune" slider. At the middle position of the slider the correction is equal to zero. At the most left position the correction makes the frequency match the note frequency, which is 1/2 tone below the note associated with the sensor. At the most right position the frequency matches the frequency of the note which is 1/2 tone above the note associated with the sensor. If the slider is moved smoothly from the most left to the most right position, the frequency of each sensor is also smoothly changed from the lowest to the highest value.

Sensor sensitivity individually set by the "Equalizer" control. "Sensitivity" control changes the sensitivity of all sensors simultaneously. The sensor sensitivity is increased when moving the "Equalizer" or "Sensitivity" slider upward.

The selectivity of sensors set by the "Selectivity" control is graphically represented in the respective window. The plot shows the sensor selectivity as a function of the audio wave frequency. The middle vertical line corresponds to the sensor own frequency. The adjacent vertical lines correspond to the frequency values differentiated by the 12-power root of two from the sensor own frequency. The frequency by the abscissa axis grows from left to right by the logarithmic scale. The selectivity plot is almost symmetric with the maximum at the own frequency.

The sensors are characterized by some inertia, displaying the rate of reaction on the appearance or disappearance of the sound in the own frequency domain. The selectivity and inertia of the sensor strictly depends one from another. The better is the sensor selectivity (the narrowest selectivity plot), the more this sensor is inert, i.e. slow. The optimum selectivity value is experimentally chosen being dependent from a particular converted melody, its, tempo, polyphony, percussion, average note duration, etc.

The values generated at the sensors' outputs are periodically scanned to detect the peaks. The peak is the sensor with the output value above the values of both adjacent sensors. The scanning is performed from left to right, i.e., from the low note sensor to the high note sensor. Before the start of each scanning cycle a uniform threshold value is set for all sensors. This threshold value depending upon the position of the "Gate" slider is presented by a horizontal dotted line in the "Spectrum" window.

When a peak is detected, the addition is made to the threshold value of the sensors presenting the notes above the current note at 12, 19 and 24 semi-tones. The addition to the threshold depends upon the peak sound intensity and the histogram shape in the "Harmonic Model" window. The histogram can be interpreted as follows: the first column presents the peak sound intensity. The second column presents the addition to the threshold value for the sensor being at 12 semi-tones away from the peak, the third column presents the same for the sensor being at 19 semi-tones away from the peak, and the fourth column – for the sensor being at 24 semi-tones away from the peak.

The harmonic model principle of the algorithm is based upon the presence of supplemental harmonics in each voice or instrumental note in addition to the main tone. Such supplemental harmonics have frequencies differentiated from the tone frequency 2-, 3-, 4-fold and more. The harmonic proportions depend upon the musical instrument or singer. The second harmonics is above the main first harmonics at 12 semi-tones exactly, the third harmonics is above the main first harmonics at 19 semi-tones with high precision, and the fourth is above the main first harmonics at 24 semi-tones exactly. The algorithm is limited by the first four harmonics, as the conversion quality is not greatly improved by further growth of the harmonics number, but the complexity is increased.

If the peak sound intensity is above the threshold value, the appropriate note generation signal is enabled. Otherwise the note disabling signal is generated, if the note was enabled.

In monophonic operation mode the scanning cycle is interrupted after the first enabled note. The previous note is disabled if different from the current note. Monophonic operation mode is set by the selection of the "Single Voice" flag in the "Method" group, and polyphone operation mode is set by the selection of the "Poliphony" flag.

The note format signal generated at this stage is filtered by the note loftiness. The filter tuning is provided in the "Filter" window (located between the piano keyboard window and "Equalizer" window). The notes with the appropriate elements on in the "Filter" window only are subjected for further processing.

The signal processed by note filtering is transposed, i.e., the note loftiness is shifted at the integer number of semi-tones, on the condition that the pre-set number of semi-tones is other from zero. The number of semi-tones for transposing is set by the "Transpose" control. At the positive control element value the notes are raised by transposing, and are lowered at the negative value.

Then the signal is directed by three separate branches: to MIDI device selected from the "Midi Out" list, to PC speaker and to built-in sequencer. The signal is applied to the MIDI device if the "MIDI" flag is marked in the "Play/Keep silence" group, to PC Speaker - if the "PC Speaker" flag is marked, and to the built-in sequencer - in the record mode, i.e., when the "Record" button is pressed.

The notes are filtered by duration when writing to the built-in sequencer. The notes shorter than the value in milliseconds set in the "Minimal Duration" window are ignored.

The resulting MIDI signal given to the built-in sequencer and to MIDI device is generated with the consideration of the selected MIDI instrument, volume and MIDI channel. The instrument is set in the "Outlet MIDI Instrument" list, the volume is set by the "Volume" control, and the channel is set in the "MIDI Channel" list.

5.0 OPERATION

The software must be tuned before conversion by setting the optimum values of all parameters during the test playing of the melody being converted. Make sure to set a flag for the device sound from which being converted in the Windows "Volume Control" program. Select the "Properties" of the "Options" menu and set the "Adjust volume for" switch to "Recording" to enable the necessary section. Set the record level of this device to normal position in this program.

First, move the "Selectivity" slider to the middle position in the AudioToMidi program. This position provides the most comfortable selectivity to start the software tuning.

Make the spectrum representation fit smoothly the "Spectrum" window by the "Sensitivity" control. The audio signal frequency characteristics should possibly be corrected by equalizer. Set the equalizer sliders' positions by pressing the mouse left button or moving the mouse cursor with pressed button. The popup menu with save/open/reset equalizer commands is enable in the equalizer window by pressing the mouse right button.

Make the explicit visual peaks in the "Spectrum" window symmetric by the "Tune" control. For example, if the guitar is tuned at 1/4 tone below the tuning fork, the slider should be moved at 1/4 control element scale to the left from the center. Proper guitar tuning is recommended, otherwise, the guitar and converted sound are not matching.

When tuning the guitar, set the "Tune" slider in the middle position. The guitar can be comfortably tuned with open strings. First, make the first harmonic peak be placed at the appropriate note level by changing the string tension. Second, make the visual peak symmetric by more precise guitar tuning. Enable the "PC Speaker" or "MIDI" option for audible control, monophonic program mode is recommended.

The next possible stage is the harmonic histogram setup in the "Harmonic Model" window. The histogram column height is set by clicking the mouse left button with the cursor at the necessary point or by moving the mouse cursor with pressed button. The popup menu with save/open/reset commands of the histogram settings is enabled in the window by pressing the mouse right button. The histogram columns should possibly match the recognized instrument

harmonics. The proportion of harmonics can be viewed in the "Spectrum" window with a single instrument note sounding and without any other sounds. If the harmonic proportions cannot be defined set the histogram similar to the default one by its shape.

Set experimentally the best note inclusion threshold by the "Gate" control. Vary the selectivity for the best conversion result. After the selectivity is changed, correct the general sensor sensitivity by the "Sensitivity" control.

Set the passed note set in the "Filter" window to reduce the number of unwanted accidental notes. A note is either included to or excluded from the set by clicking the mouse left button with the cursor at the necessary point or by moving the cursor with the mouse button pressed. This can be done to all of the same name notes by double-click of the mouse left button in the note area. Another way to do that is to click the mouse right button at the selected note key in the piano keyboard window. The popup menu with save/open/reset commands of the filter settings is enabled in the "Filter" window by pressing the mouse right button. The cleaning command for the passed note set is also available in the menu.

The conversion quality can be improved by narrowing the interval of passed notes by the filter tuning. As a rule, the percussion noise is concentrated in the low frequency domain. Consequently, if the musical piece has percussion instruments, the lower interval should be cut. In many cases the melody to be converted is limited by two or three octaves, and in rare cases it is limited by four or more octaves. Hence, the upper interval of passed notes can also be cut.

The musical piece being converted may have a certain key. In this case the passed note set must have only the notes associated with the pre-set key with possible addition of raised or lowered key degrees. The software package includes the ready-made files of filter settings for all possible major and minor keys. The files for minor keys with the added raised seventh degree are also included. The file "Joe Dassin (+_Bm7).mid" obtained by the melody conversion in B minor key with the added raised seventh degree, i.e., A sharp, is enclosed as an example.

The conversion result can be listened to in real time if the "MIDI" option of "Play/Keep silence" group is on. Monophonic melody result can be listened to by the built-in speaker, if the "PC Speaker" is on.

Select the wanted MIDI instrument from the "Outlet MIDI Instrument" list for final MIDI signal. The wanted volume is set in the window "Volume". The number of the shift semi-tones is set in the window "Transpose". If several sound tracks each intended for a particular MIDI instrument are used for generation, a unique MIDI channel must be set for each track. Select a channel from the "MIDI Channel" list. Note: tenth channel is allocated for the percussion instruments and sound effects.

Select the optimum value of the minimum note duration in milliseconds in the "Minimal Duration" window.

Press the button "Record" to record the track in the built-in sequencer. Press the button "Pause" or start recording from playing station to listen to the tracks with a new track being imposed when recording.

6.0 LICENSE

This software product is a freeware version. The author is not responsible neither for possible errors related to the operation of this software product nor for the consequences of such errors. The author rights to this software product are the property of Alexey Egorov. This software product cannot be sold, commercialized or distributed in a modified. This software product can be distributed together with the present documentation.

7.0 CONTACT

Please find the information and files related to this product at Error: Reference source not found. Please mail your suggestions and tell about the bugs to Alexey Egorov Error: Reference source not found.

Alexey Egorov,
Error: Reference source not found
Error: Reference source not found

September 7, 1999