

ДИНАМИЧЕСКИ РЕКОНФИГУРИРУЕМЫЙ ПРОЦЕССОР АДАПТИВНОГО ПАКЕТНОГО ВЭЙВЛЕТНОГО ПРЕОБРАЗОВАНИЯ ДЛЯ СИСТЕМ КОДИРОВАНИЯ АУДИО СИГНАЛОВ

Петровский А.А.

Белорусский государственный университет информатики и радиоэлектроники,
ул. П.Бровки 6, Минск, Беларусь, e-mail: apetrows@bas-net.by

Реферат. В докладе предлагается структура динамически реконфигурируемой системы для кодирования аудио сигналов на основе цифрового процессора обработки сигналов (host-процессор) и реконфигурируемого процессора на FPGA (дискретное пакетное вэйвлетное преобразование (ДПВП) на распределенной арифметике). Поиск лучшей декомпозиции ДПВП осуществляется на основе алгоритма, базирующегося на вычислении перцептуальной энтропии как меры качества декомпозиции.

Введение

Стремительное развитие систем мультимедиа и интернет обусловили актуальность исследований и разработку новых методов высококачественной компрессии аудио сигналов [1] и аппаратных платформ для них [2,3]. Известно, что в CD технологии частота дискретизации аудио сигналов равна 44,1 кГц, а разрядность АЦП – 16 бит. Следовательно, скорость передачи последовательного кода будет составлять 705,6 Кбит/сек., что достаточно много для стандартного типа мультимедиа приложений. Для обработки нестационарных сигналов применяется новое теоретическое направление цифровой обработки сигналов – дискретное вэйвлетное преобразование, которое в отличие от известных, позволяет осуществлять обработку таких сигналов в частотной и временной областях, проводить анализ сигналов с переменным частотным и временным разрешением. На основе данного преобразования возможна организация адаптивного дискретного пакета вэйвлетного преобразования (АДПВП), на базе которого могут быть реализованы системы компрессии аудио сигналов.

Построение адаптивной системы компрессии требует значительных вычислительных затрат. Доминирующее положение, для решения задачи компрессии аудио сигналов, занимают специализированные устройства, базирующиеся на двух или более процессорах цифровой обработки сигналов (ЦПОС). Однако, более прогрессивным направлением, решающим данную проблему в полном объеме, заключается в применении ЦПОС и программируемых логических матриц (в частности FPGA - Field Programmable Gate Arrays), как единого ядра системы [5].

Целью данного доклада является показать структуру динамически реконфигурируемого процессора для систем кодирования аудио сигналов на основе алгоритма адаптивного ДПВП базирующегося на вычислении перцептуальной энтропии как меры качества декомпозиции.

Структура аудио кодера

В данном разделе будет описана структура кодера, применяемого для кодирования моно аудио сигналов с частотой дискретизации 44.1 кГц.

Общая структура кодера аудио сигналов показана на рис.1.

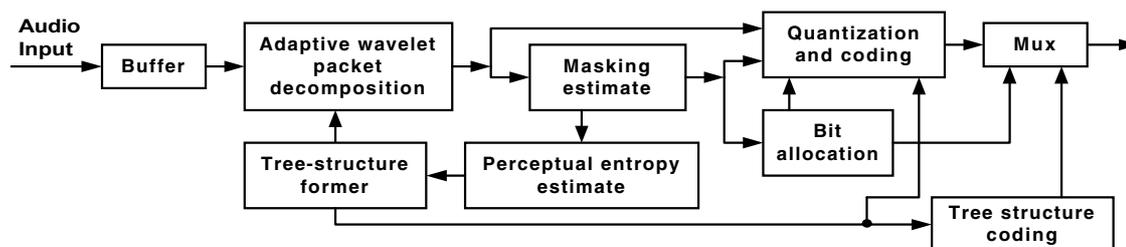


Рис. 1 Структура кодера

Характеристика структуры кодера следующая:

1. Анализ входного сигнала выполняется банком фильтров, реализованного на базе ДПВП адаптированного к источнику (аудио сигнал);
2. Выполняется сегментация входного сигнала на фреймы длиной 256 отсчетов;
3. Оценка маскирующих порогов (Masking estimate) осуществляется в вэйвлетной области [6], а не в частотной области [7], что позволяет уменьшить вычислительную сложность алгоритма;
4. Формирование динамически реконфигурируемой структуры дерева ДПВП выполняется на основе вычисления перцептуальной энтропии (Perceptual entropy estimate) для каждой ветви дерева ДПВП;
5. Нахождение субоптимальной структуры дерева (Tree structure former) для каждого сегмента входного сигнала, что исключает избыточность вычислений в блоке (Adaptive wavelet packet decomposition);
6. Квантование вэйвлетных коэффициентов (Quantization) выполняется линейным квантователем для ветвей одного уровня и соответственно количество квантователей, применяемых в структуре кодера, зависит от глубины построенного дерева ДПВП;
7. Кодирование (Coding и Tree structure coding) и декодирование (Decoding и Wavelet packet tree computation) коэффициентов преобразования и дополнительной информации (структура дерева ДПВП, длина

кодированной последовательности и т.д.) выполняется алгоритмом Хаффмана (Huffman) на базе экспериментальных гистограмм распределения вероятности.

8. Структура декодера реализована на основе синтезирующего банка фильтров инверсного ДПВП.

Структура декодера аудио сигналов показана на рис.2.

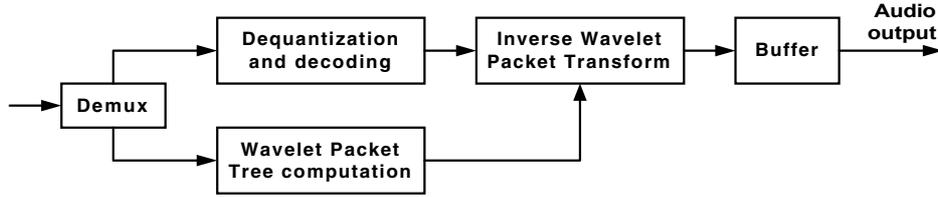


Рис. 2 Структура декодера аудио сигналов

Адаптивное ПДВП (АПДВП)

Ядром кодера и декодера является ДПВП (обобщение диадического вэйвлетного преобразования), которое позволяет получать богатое множество структур, путем соответствующих его декомпозиций. Структура ПДВП больше согласуется со свойствами сигнала, чем вэйвлетное преобразование. Причем, ПДВП ассоциируется с алгоритмом выбора лучшей структуры \mathbf{p} преобразователя из множества структур, путем измерения и минимизации определенной меры качества \mathbf{J} , зависящей от обрабатываемых данных. Например, в [8] число ненулевых коэффициентов преобразования после сравнения с неким порогом используется в качестве такой меры.

Критерии оптимизации, основанные на энтропии, хорошо соответствуют процедуре выбора эффективного бинарного дерева ПДВП для приложений кодирования аудио сигналов [4], потому что наиболее адекватно описывают информационные свойства представления заданного сигнала. При этом, для перцептуального кодирования аудио сигналов нужна такая декомпозиция ПДВП, при которой минимизируется скорость передачи и обеспечивается высокое качество восприятия человеком декодированного сигнала [4]. Следовательно, банк вэйвлетных фильтров $(\mathbf{p}_\infty, \mathbf{z}_\infty, \mathbf{d}_\infty)$, согласованный с критической шкалой частот восприятия акустической информации человеком, есть предельная структура для метода перцептуального кодирования [9].

Конструкция алгоритма выбора лучшей декомпозиции \mathbf{p} ПДВП заключается в вычислении меры качества (целевой функции) \mathbf{J} для каждого узла \mathbf{z} дерева ПДВП глубины \mathbf{d} . Затем у дерева отсекаются ветви согласно условию **если** $\mathbf{J}_{\text{cur}} > \mathbf{J}_{\text{child}}$, **то** $\mathbf{J}_{\text{cur}} = \mathbf{J}_{\text{child}}$, **иначе** отсечь подчиненные узлы, где \mathbf{J}_{cur} - мера качества для текущего узла дерева ПДВП, а $\mathbf{J}_{\text{child}}$ - сумма мер качества соответствующих подчиненных узлов. Результирующие коэффициенты ПДВП будут соответствовать лучшей структуре дерева \mathbf{p} ПДВП после отсека ветвей.

Динамическая реконфигурация ПДВП

Функция Джонсона [10], которая зависит от перцептуальной энтропии, выбирается мерой качества декомпозиции ПДВП, и характеризует минимальное число бит на отсчет сигнала, необходимого для получения, так называемого “прозрачного” кодирования аудио сигнала, при котором шумы квантования не воспринимаются слухом человека.

Среди множества целевых функций \mathbf{J} выделим функции, которые аддитивны к последовательности данных \mathbf{x} :

$$\mathbf{J}^{\text{add}}(0) = 0, \mathbf{J}^{\text{add}}(\mathbf{x}) = \sum_i \mathbf{J}^{\text{add}}(\mathbf{x}(i)).$$

Для узла \mathbf{z} дерева ПДВП глубины преобразования \mathbf{d} предлагается следующая целевая функция:

$$\mathbf{J}(\mathbf{z}) = \sum_{\forall X(\mathbf{z})} \log_2(\mathbf{SMR}(\mathbf{z})),$$

где \mathbf{SMR} - отношение коэффициентов ПДВП $X(\mathbf{z})$ к порогу восприятия $T(\mathbf{z})$ в узле \mathbf{z} , как

$$\mathbf{SMR} = \left\lceil \frac{|X(\mathbf{z})|}{\sqrt{c \cdot T(\mathbf{z})/K(\mathbf{z})}} \right\rceil,$$

здесь $K(\mathbf{z})$ – количество коэффициентов ПДВП в узле \mathbf{z} , $\lceil \mathbf{a} \rceil$ – округление \mathbf{a} до ближайшего целого значения с целью избежания переполнения разрядной сетки вычислителя, коэффициент c обусловлен выбором типа квантователя, $T(\mathbf{z})$ - порог восприятия в узле \mathbf{z} , вычисляемый по методике [9].

Алгоритм декомпозиции ДПВП для поиска лучшей структуры:

1. Начало. Для корневого узла вычисляется начальное значение функции меры $\mathbf{J}(\mathbf{z}_i)$, где $\mathbf{z}_i = \mathbf{x}$ - фрэйм входного аудио сигнала. Объявляется первоначальная структура $\mathbf{p}(\mathbf{z}_i)$.
2. Осуществляется декомпозиция сигнала с использованием банка фильтров без потерь на основе ортогональных вэйвлетных фильтров по структуре $\mathbf{p}(\mathbf{z}_i)$.

3. Вычисляется целевая функция для узлов декомпозиции ($J(z_j)$ и $J(z_k)$).
4. Если $J(z_i) \geq J(z_j) + J(z_k)$, тогда подчиненные узлы становятся текущими $z_i = \{z_j, z_k\}$ и структура $p(z_i) = p(z_j) + \{p(z_k)\}$;
 если $p < p_\infty$ и $z_i < z_\infty$, тогда возврат к п.2 иначе переход к п.5.
 иначе узлы z_j и z_k отсекаются и возврат к п.2.
5. Конец. Коэффициенты лучшего ДПВП соответственно представлены структурой $p(z_i)$.

Таким образом, для каждого фрейма аудио сигнала узел i разделяется на два узла j и k , если и только если сумма весовых функций, определенных в узлах j и k , меньше, чем та, что в узле i . Данный алгоритм позволяет получать субоптимальные решения. На рис. 3 показана субоптимальная структура банка вэйвлетных фильтров для одного из фреймов аудио сигнала.

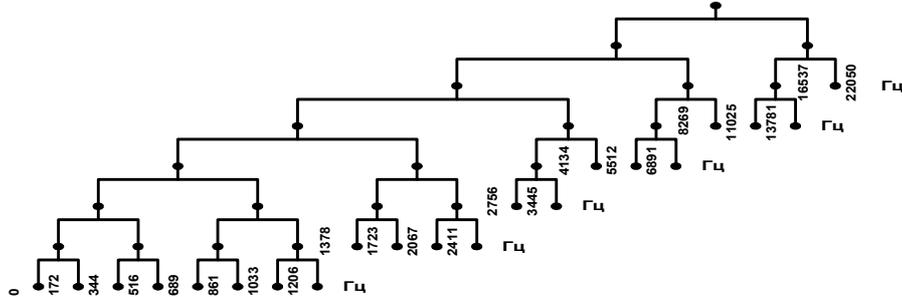


Рис. 3 Пример субоптимальной структуры банка фильтров

Реконфигурируемая система

Реконфигурируемая система для выполнения операций кодирования и декодирования аудио сигналов в реальном времени построена на базе FPGA и цифрового процессора обработки сигналов. Задача вычисления адаптивного банка фильтров (анализ и синтез) возлагается на реконфигурируемую часть системы (FPGA + распределенная арифметика). Базовый элемент ДПВП, представленный банком фильтров, выполнен на распределенной арифметике (рис.4). Здесь входные отсчеты $x(k)$, которым соответствует битовая последовательность B_x , обрабатываются фильтрами нижних и верхних частот с конечной импульсной характеристикой. Результатом вычислений после децимации на 2 есть две выходные битовые последовательности B_v отсчетов $v(k)$ и B_w - $w(k)$. В соответствии со структурой дерева ДПВП количество ветвей с каждой новой ступенью возрастает вдвое, а поток данных в ветви уменьшается на такую же величину, то для обеспечения $\approx 100\%$ загрузки модулей фильтрации вводится блок буферизации и переключателей (Buffer/SW) (рис.5). Сигнал SYNC синхронизирует запись промежуточных результатов $v_i(k)$ и $w_i(k)$ и чтение исходных данных для последующего уровня $x_i(k)$ из временных буферов (BF_L и BF_H) модулей $1 \dots L$ и выборку сформированных результатов (OUTPUT). Схема вычисления ДПВП на FPGA показана на рис.6, согласно которой для каждого уровня дерева ДПВП применяется один блок фильтрации (Filter bank) и буферизации/переключателей (Buffer/SW) за исключением последних ветвей.

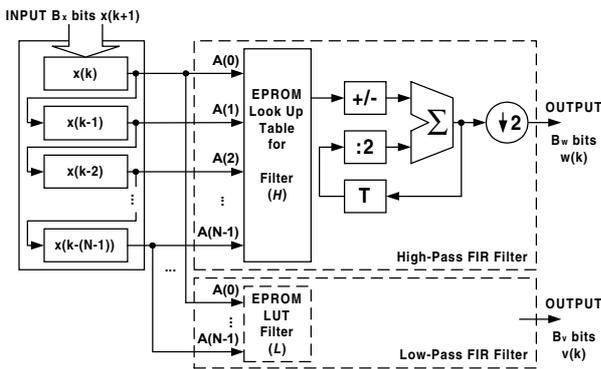


Рис. 4 Анализирующий/синтезирующий банк фильтров на распределенной арифметике.

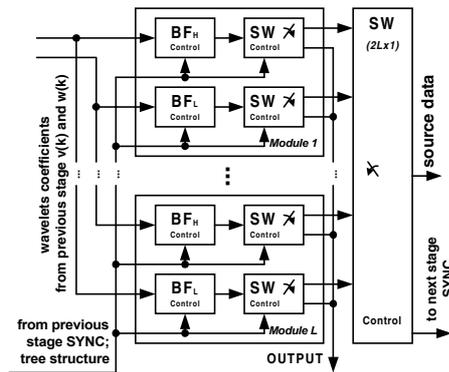


Рис. 5. Общая схема блока буферизации и переключателей

Изменение структуры FPGA осуществляется в соответствии с алгоритмом задачи, вычисление маскирующих порогов, кодирование и декодирование и операции общего назначения выполняются на ЦПОС

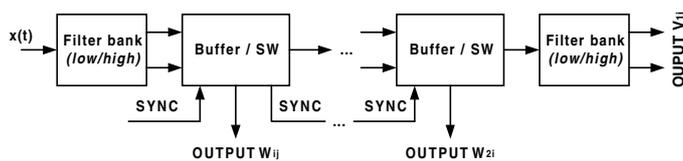


Рис. 5 Схема вычисления ПДВП на FPGA.

Реконфигурируемая система, объединяющая ЦПОС и FPGA, обладает рядом преимуществ перед другими альтернативами: быстрее и меньше, чем системы общего назначения; низкая стоимость разработки, чем специализированных систем; различные структуры систем могут быть реализованы без аппаратного изменения платформы.

Литература

1. Painter T., Spanias A. A review of algorithms for perceptual coding of digital audio signals // The proc. of the intern. conference DSP97, Creece 1997. – pp. 179-208.
2. Jayant N.S., Chen E.Y. Audio compression: technology and applications. // AT&T technical journal, vol.74, №2, 1995. – pp. 23-34.
3. Petrovsky A.I.A. Multi-DSP based hardware platform for wavelets and filter banks // Proc. of the International Workshops on System, Signal and Image Processing, June 1999, Bratislava, Slovakia, - pp. 111-114.
4. Jayant N., Johnston J., Safranek R. Signal compression based on models of human perception. – Proc. of the IEEE, vol. 81, № 10, Oct., 1993. - pp. 1385–1422.
5. Bielawski K., Petrovsky A.I.A. Dynamic non-uniform filter bank constructing algorithms for reconfigurable speech processing system based on the FPGA-device and TMS320C31 // The 3rd European DSP education and research conference, France, 20-21 Sept. 2000. – 8p.
6. Ал.А. Петровский, Расчет маскирующих порогов для аудио кодеров на базе пакетного дискретного вэйвлетного преобразования // Радиотехника и электроника, вып. 25, Мн.: БГУИР, 2000. – с. 34-44.
7. N.Ruiz Reyes, M.Rosa Zurera, F. Lopez Ferreras, and et al. A new cost function to select the wavelet decomposition for audio compression // IEEE Nordic Signal processing symposium, NORSIG'2000, Sweden, June 13-15, 2000. – pp.331-334
8. Coifman R., Meyer Y., Quake S. and Wickerhauser V. Signal processing and compression with wave packets. - In the Numerical Algorithms research group, New Haven, CT: Yale University, 1990.
9. Zwicker E., Fastl H., Psychoacoustics: Facts and Models. - Springer-Verlag, 1990
10. Johnston J.D., Transform coding of audio signals using perceptual noise criteria, IEEE Trans. On Select., Areas Commun., vol. 6, Feb. 1988, pp. 314-323.



DYNAMIC RECONFIGURABLE PROCESSOR ADAPTIVE WAVELET PACKET DECOMPOSITION FOR AUDIO SIGNAL CODER

Petrovsky A.A.

Computer Science Department,
Belarusian Sate University of Informatics and Radioelectronics,
6, P.Brovky st., Minsk 220027, Republic of Belarus, e-mail: apetrows@bas-net.by

Audio encoder structure

The general audio encoder structure is shown in figure 1.

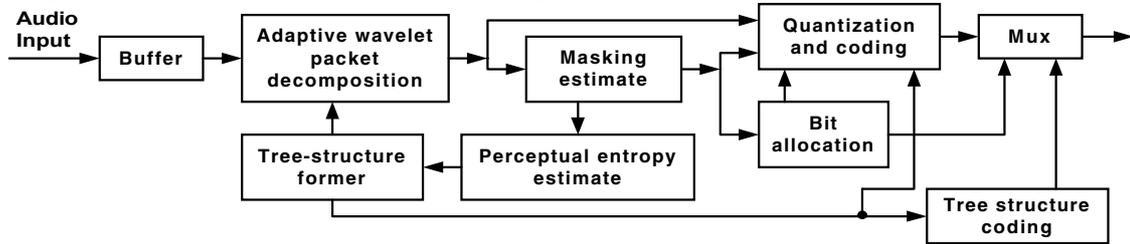


Figure. 1 Encoder structure

The main features of the audio coder are the following:

1. The input signals are segmented in a frames of 256 samples and analyzed with a time-varying filter bank that implements a wavelet packet decomposition adapted to the source (audio signals);
2. The masking thresholds are estimated in a wavelet domain [6]. Using other transformation algorithms (DFT, DCT) is not necessary. In this case, the translation errors from time/frequency to wavelets domain in the psychoacoustic information are not present;
3. Dynamic reconfigurable wavelet packet (WP) tree structure is formed basing on perceptual entropy calculations for each tree nodes of WP. Finding suboptimal WP tree structure in the block (Tree structure former) for each input segments eliminates computational redundancy in WP;
4. WP coefficients are quantized using uniform quantizer for each current level tree-nodes and number of quantization elements depends on the WP tree depth;
5. Decomposition coefficients and side information (the WP tree structure, the length of encoded sequence and etc.) are coded using Huffman codebook based on experimental histograms;
6. At the decoder, the implemented synthesis filter bank based on inverse WP transform reconstructs the input sequence.

The decoder structure is shown in figure 2.

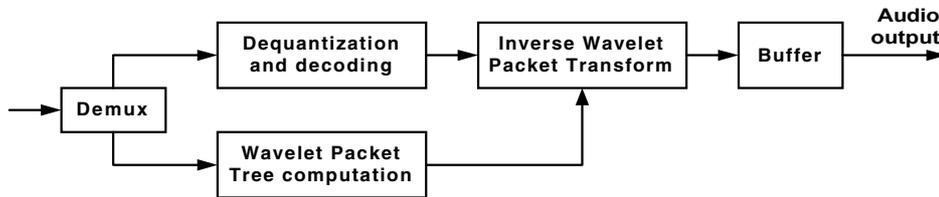


Figure. 2 Decoder structure.

Dynamic reconfigurable WP decomposition

Johnson function [10], which depends on perceptual entropy, is selected as WP decomposition cost function and defines the minimal number of bits per sample, that is needed to obtain the transparent audio coding, where quantization noise is not audible for humans.

The cost function for the node \mathbf{z} of WP tree depth \mathbf{d} is determined as

$$\mathbf{J}(\mathbf{z}) = \sum_{\forall X(\mathbf{z})} \log_2(\mathbf{SMR}(\mathbf{z})),$$

where \mathbf{SMR} is a relation between WP coefficients $X(\mathbf{z})$ and perceptual threshold $T(\mathbf{z})$ in the node \mathbf{z} , as

$$\mathbf{SMR} = \left\lceil \left| \frac{X(\mathbf{z})}{\sqrt{c \cdot T(\mathbf{z}) / K(\mathbf{z})}} \right| \right\rceil,$$

where $K(\mathbf{z})$ is number of the WP coefficients in node \mathbf{z} , $\lceil \mathbf{a} \rceil$ is rounding \mathbf{a} to the nearest integer, it is needed to escape the overflow, coefficient c is defined by quantizer type, $T(\mathbf{z})$ is perceptual threshold in the node \mathbf{z} calculated according to the method [9].

Algorithm: Best WP tree structure searching:

1. Begin. For the root node is calculated the initial value of the cost function $\mathbf{J}(\mathbf{z}_i)$, where $\mathbf{z}_i = \mathbf{x}$ is the first input audio signal frame. The initial WP tree structure $\mathbf{p}(\mathbf{z}_i)$ is defined.
2. WP decomposition based on orthogonal wavelets function according to the structure $\mathbf{p}(\mathbf{z}_i)$ is calculated.

3. The cost functions for the decomposition nodes ($J(z_j)$ и $J(z_k)$) are calculated.
4. If $J(z_i) \geq J(z_j) + J(z_k)$, then the children nodes become a current nodes $z_i = \{z_j, z_k\}$ and structure $p(z_i) = p(z_j) + \{p(z_j), p(z_k)\}$;
 if $p < p_\infty$ and $z_i < z_\infty$, then return to p.2 else go to p.5.
 else nodes z_j and z_k are cut off and returning to p.2.
5. End. The resulting structure $p(z_i)$ represents the best WP decomposition.

In that way, for each frame of the input audio signal node i is split on two nodes j and k , if and only if the sum of the cost functions in the nodes j and k less then in the node i . Such algorithm permits getting suboptimal solutions.

Reconfigurable hardware

The novel reconfigurable system based on FPGA and digital signal processor (DSP) for audio coding and decoding in real-time is being described. Analysis/synthesis filter bank is organized as adaptive WP decomposition and implemented on the reconfigurable hardware platform FPGA using distributed arithmetic (DA). The basic decomposition of WP expressed as filter bank designed on DA is suggested in figure 3. The input samples $x(k)$, represented as bit stream B_x , are filtered with high and low FIR filters. Two output sequences B_v of the sample $v(k)$ and B_w of the $w(k)$ are filtering result after decimation by factor 2. In according with a WP tree structure, the number of nodes is increased in two times but the data stream is decreased to the same degree. For providing $\approx 100\%$ working time utilization the block Buffer/Switch (figure 4) was introduced. Signal SYNC is synchronized writing intermediate data $v_i(k)$, $w_i(k)$ and reading initial data for the next stage $x_i(k)$ from temporary buffers (BF_L и BF_H) of module $L..L$ and extracting results (OUTPUT). Figure 5 shows the schema of computation WP decomposition implemented on FPGA, where only one filter bank and one block Buffer/Switch are needed to implement each level of WP except for the last one.

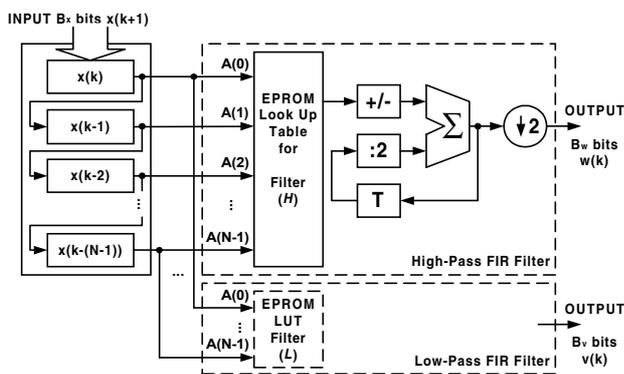


Figure 3 Analysis/Synthesis filter bank based on DA

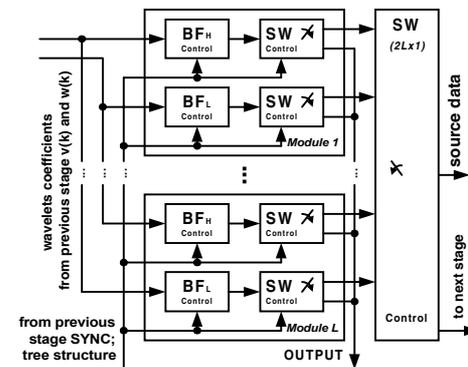


Figure 4 Buffer/Switch block

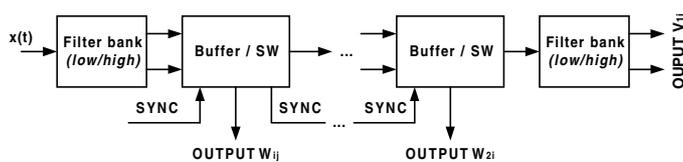


Figure 5 Computation WP decomposition on FPGA

Control signals, FPGA reprogramming signals in accordance with process, masking thresholds calculation, encoding and decoding and etc. are formed and generated by DSP.

That reconfigurable system uniting digital signal processor and FPGA, has a lot of advantages in comparison with other alternatives: quicker and smaller as general purpose systems; low development cost compared with specific systems, different system structures can be realized without hardware changes.