Wavelet Transform for Analyzing Fog Visibility

Christoph Busch and Eric Debes, Computer Graphics Center, Germany

THE AUTHORS PRESENT A NEW TECHNIQUE FOR ESTIMATING THE DISTANCE OF VISIBILITY IN FOG CONDITIONS. BASED ON A PSYCHOVISUAL MODEL AND ON CONTRAST ESTIMATION WITH WAVELET TRANSFORM, THEIR TECHNIQUE FARED WELL WHEN COMPARED TO A DIRECT APPROACH BASED ON LOCAL CONTRAST CALCULATION.

A GROWING NUMBER OF COMputer-vision systems have emerged for traffic analysis and control. Increased traffic, along with steady efforts to improve driver security, have encouraged the development of sophisticated video-based systems, providing statistical measures of traffic flow for actual traffic information on variable message signs. Investigations show that VMS reduce traffic jams, thus minimizing the burden placed on financial and environmental resources. Moreover, traffic control systems reduce the number of accidents to 45%.[1]

Of particular interest to drivers is information on low visibility in fog—a dangerous driving condition that causes many accidents. Because places where fog frequently rolls in are often well known, message signs inform motorists of visibility distance. Traditionally, the visibility distance is estimated with transmissiometers, but we suggest replacing conventional fog-detection sensors with the integrated functionality that a computer-vision system can provide.

Our method, based on a B-spline wavelet transform, analyzes video frames used in traffic-control systems and derives a visibility-distance measure with extended signal-analysis techniques. With our wavelet-based approach, we can inform drivers in real time of the presence of fog through VMS located a few kilometers before the foggy area. Analyzing the entire traffic scene globally, our method provides more precise input to fog information systems than the traditional transmissiometers that analyze fog density locally beside the roadway.

Wavelet-based contrast estimation

Because our approach is based on psychovisual perception, we briefly introduce the basic notion of psychovisual perception in fog and explain the relation between contrast, gradient, and wavelets.

Psychovisual perception in fog. Visual perception is a rather complex mechanism, because it is based on space, time, and frequency analysis. The visual system receives a signal, and the brain then codes the signal's different characteristics: intensity, color, shape (contrast and direction), and movement (direction and speed). To see an object, the object must stand out from its background, which is why shape perception depends on the contrast between the object and its background. The contrast is usually defined as

$$C = \frac{|I_o - I_b|}{I_o + I_b}, \qquad (1)$$

where I_o is the object's intensity and I_b is the background's intensity.

In 1948, S. Duntley formulated a law indicating that in fog conditions, contrast decreases exponentially with distance:

$$C = C_o e^{-Kd}, \qquad (2)$$

where C is the contrast the visual system receives, C_o is the real contrast, d is the distance, and K is an attenuation coefficient that characterizes the fog conditions.

As the origin of all psychovisual effects, contrast also plays an important role in the perception of movement, speed, and direction. The visual system cannot correctly estimate an object's speed with a low contrast,[2] and fog causes errors in the estimation of an object's shape, distance, and speed.[3] Furthermore, the measure of contrast is independent of daylight illumination.[4]

Contrast, gradient, edges, and wavelets. To determine visibility in fog, we must quantify the contrast and compare this value to a reference value operating as a threshold. There is a high correlation between the contrast of different gray-level areas and the gradient's value in the region where we measure the contrast. A low contrast corresponds to a low gradient and vice versa, because contrast and gradient both characterize local variations in gray levels.

In fact, what we see in fog is only high gradients in the different gray levels. When driving in fog, the visual system only utilizes the parts of the image characterized by high contrast (or high gradient). Thus, to determine if a pixel is visible, we just calculate the gradient. If the gradient in this pixel is higher than a threshold corresponding to a 5% contrast, it indicates that the driver can see the pixel. The CIE (International Commission on Lighting) defined the 5% contrast threshold as the sufficient contrast for a human eye to see the difference between two gray levels.

There is a close correlation between gradient-measure and edge-detection techniques. In particular, we can use multiscale edge detection to calculate the gradient in 2D signals. For a complete description of the edge-detection technique, see Stephane Mallat and Wen Lian Hwang's "Singularity Detection and Processing with Wavelets."[5] Mallat and Hwang explain how the wavelet transform can express the gradient and how modulus M of the gradient is proportional to

$$Mf(x,y) = \sqrt{\left|W^1 f(x,y)\right|^2 + \left|W^2 f(x,y)\right|^2} \quad (3)$$

In Mallat's edge-detection theory, angle α of the gradient is also taken into account, so that the edge points are located in points where modulus M has a local maximum in direction α. However, our aim is to calculate the gradient, not to detect edges, so we don't need to calculate the angle. We simply calculate the gradient for each point of our image and, if this gradient is higher than a value corresponding to a 5% contrast, the pixel is visible regardless of the corresponding edge's orientation.

B-splines-based contrast estimation

Here we explain why the properties of B-Spline wavelets are interesting for edge-detection purposes, and we present B-splines with specific small support—especially well suited for visibility analysis.

> ***The goal is to determine the distance of visibility in fog conditions from images taken by a video camera installed overhead in a traffic-lane motorway.***

The Gaussian kernel and B-splines. The central problem involves choosing the proper smoothing function $\theta(x, y)$. For edge detection, we need a function that is well localized in the frequency and spatial domain; in particular, the variance Δx of the filter, as well as the variance $\Delta\omega$ of the filter's spectrum, should be small. It is well known that these two localization requirements conflict. David Marr and Ellen Hildreth have used the Gaussian kernel for edge detection,[6] because this function optimizes the Heisenberg uncertainty principle, which states that

$$\Delta x \Delta\omega \geq \frac{1}{4}\pi \quad (4)$$

Another reason why the Gaussian kernel is often used in edge detection is that the human retina's response resembles a Gaussian function. There are receptive field profiles in the mammalian retina and visual cortex, and superposition of Gaussian derivatives effectively models the measured-response profiles.[7]

In practice, because the computational load becomes extremely high with the Gaussian kernel when the scales grow, many techniques have been proposed for efficient implementation of scale-space filtering in computer vision. Tomaso Poggio and his colleagues, for example, used B-splines to approximate the Gaussian kernel with efficient implementation.[8]

The advantages of B-splines, in comparison with the Gaussian kernel, are that these functions lead to fast implementations and are still a close approximation to the Gaussian kernel. Both the B-splines and their Fourier transforms converge to the Gaussian function as the order of the spline tends to infinity:[9]

$$\beta^n(x) \approx \sqrt{\frac{6}{\pi(n+1)}} \exp\left(-\frac{6x^2}{n+1}\right) \quad (5)$$

Experiments show that the cubic B-spline is already nearly optimal in terms of time and frequency localization in the sense that its variance is within 2% of the limit that the uncertainty principle specifies.

Both physiological and biological experiments have shown that a Gaussian kernel can model the human visual system. Therefore, the B-spline is also suitable for modeling biological vision due to its similarity to Gaussian functions. However, because we can expect a fast implementation using the B-splines approach, it is preferable to select the B-spline rather than the Gaussian function to calculate the contrast in an image with a good model (see the "B-splines for edge detection" sidebar).

The wavelet-based method vs. direct calculation

Our goal is to determine the distance of visibility in fog conditions from images taken by a video camera installed overhead in a traffic-lane motorway. To determine the contrast, we used two methods—one based on direct contrast calculation and the other based on the edge-detection theory. Figure 1 shows two images taken with a video camera, which we use to present the results for the two methods. The image in Figure 1a was taken in a clear atmosphere with good visibility, and Figure 1b was taken in fog conditions with a visibility distance of about 100 m.

Direct contrast calculation. The first method aims to determine the distance of visibility directly, on the raw data provided by the video sensor. The contrast quantification is calculated directly on a mask of a few pixels.

We can easily extend the results for the 1D case to 2D. In this case, the smoothing function is defined as

$$\beta^n(x,y) = \beta^n(x)\beta^n(y) \tag{13}$$

and the 2D wavelet transform is given by:

$$W_{2^j} f(x,y) = \begin{bmatrix} W^1_{2^j} f(x,y) \\ W^2_{2^j} f(x,y) \end{bmatrix} = \begin{bmatrix} f * \psi^{n,1}_{2^j}(x,y) \\ f * \psi^{n,2}_{2^j}(x,y) \end{bmatrix}, \tag{14}$$

where

$$\psi^{n,1}(x,y) = \psi^n(x)\beta^n(y), \psi^{n,2}(x,y) = \psi^n(y)\beta^n(x) \tag{15}$$

and $\psi^n(x)$ is the same wavelet as in the 1D case defined in Equation 1. With Equation 14,

$$W_{2^j} f(x,y) = 2^j \vec{\nabla}(f * \beta^n_{2^j})(x,y) = 2^j \begin{bmatrix} \frac{\partial}{\partial x}(f * \beta^n_{2^j})(x,y) \\ \frac{\partial}{\partial y}(f * \beta^n_{2^j})(x,y) \end{bmatrix} \tag{16}$$

we get the partial derivation and thus the gradients according to the main axes. We can extend Equation 8 in the "Theory of B-Splines" sidebar and Equation 6 to the 2D case:

$$\hat{\beta}^n(2\omega_x, 2\omega_y) = H(\omega_x)H(\omega_y)\hat{\beta}^n(\omega_x, \omega_y) \tag{17}$$

$$\hat{\psi}^{n,1}(2\omega_x, 2\omega_y) = G(\omega_x)\hat{\beta}^n(\omega_x, \omega_y) \tag{18}$$

$$\hat{\psi}^{n,2}(2\omega_x, 2\omega_y) = G(\omega_y)\hat{\beta}^n(\omega_x, \omega_y) \tag{19}$$

Extending Equation 12 we formulate the smoothing operator:

$$S_{2^j} f(x,y) = f * \beta_{2^j}(x,y), B_{2^j}(x,y) = \frac{1}{2^j}\beta\left(\frac{x}{2^j}, \frac{y}{2^j}\right) \tag{20}$$

so that the fast recursive algorithm can be expressed as

$$\begin{cases} S_{2^j} f = S_{2^{j-1}} f * (h,d)_{\uparrow 2^{j-1}} \\ W^1_{2^j} f = S_{2^{j-1}} f * (d,g)_{\uparrow 2^{j-1}} \\ W^2_{2^j} f = S_{2^{j-1}} f * (g,d)_{\uparrow 2^{j-1}} \end{cases} \tag{21}$$

where $I * (h, g)_{\uparrow 2^{j-1}}$ is the convolution of the lines and columns of the image signal with the 1D filter $[h]_{\uparrow 2^{j-1}}$ and $[g]_{\uparrow 2^{j-1}}$ respectively. Sign d stands for the Dirac filter, whose impulse is equal to 1 for $n = 0$ and 0 otherwise.

With this method, the two wavelet-transform components are proportional to the two components of the gradient

$$\vec{\nabla}(f * \beta^n_{2^j}).$$

For each scale 2^j, the modulus of the gradient vector is proportional to

$$M_{2^j} f(x,y) = \sqrt{\left|W^1_{2^j}\right|^2 + \left|W^2_{2^j}\right|^2}. \tag{22}$$

Each pixel for which the gradient is higher than a threshold value is visible through fog. To determine the threshold value corresponding to a contrast of 5%, we calculate the wavelet transform on a test image with a contrast of 5%. We take as the threshold the highest value of the resulting transform, which we find at the position of the edge between the two different gray-level areas.

Reference

1. M. Unser, A. Aldroubi, and M. Eden, "B-Spline Signal Processing: Part I and II," *IEEE Trans. Signal Processing*, Vol. 41, No. 2, Feb. 1993.

for three types of wavelets with different support lengths:

- the splines from Michael Unser, with an FIR filter with 27 coefficients,
- the splines given by Mallat in his work on multiscale edges,[10] and
- the splines we calculated in the "B-splines for edge detection" sidebar.

Figure 3 shows the wavelet transform in the first iteration. It's obvious that these wavelets are too large for our application. This comes from the fact that B-spline wavelet functions are convoluted with the image; they take into account the variations located in neighboring pixels in their contrast measurement. For our application, we need wavelet functions with a smaller support, as in the case for Mallat's wavelet function (see Figure 4).

The wavelets that we calculated (in the "B-splines for edge detection" sidebar) have the advantage of introducing almost no blurring in the resulting images, because only the pixels where the variation occur are taken into account when calculating the contrast. These wavelets fit very well with our application, because we only want to detect variations between two or three neighboring pixels. We show the resulting images for this type of wavelet in Figure 5.

The visibility distance is calculated with a transformation from image-to-world coordinates. If this distance is higher than 1,000 m (such as in Figure 5a), according to the definition of fog given by the German federal traffic department, there is no fog. We round visibility between 1,000 m and 300 m to the nearest multiple of 50 m, and a visibility distance smaller than 300 m to the nearest multiple of 10 m. For example, the distance of visibility is 120 m for Figure 5b. Our system's accuracy is much better than what we need to give such results, because the precision depends on the distance in the real world between two image pixels. For a distance between 1,000 m and 300 m, this distance is much smaller than 50 m, and for a distance between 300 m and 50 m, it is much smaller than 10 m. According to these requirements and specifications, our system gives strong results for this application.

DUE TO ITS PSYCHOPHYSIOLOGical foundation, Mallat's edge-detection theory is well suited to estimate gradients in an image and, hence, to describe visibility in fog situations. To overcome the drawbacks that arise with limited video-frame resolution, we present B-splines with specific small support capable of fog detection. Even though the accuracy of the two compared methods differs only slightly, the wavelet-based approach has a major advantage in its close similarity to the driver's perceptual system, because it is based on Mallat's vision model. This theory is much closer to the human visual system than the contrast theory. Indeed, it takes into account not only the direct difference of intensities between objects, but also the intensity gradient in the

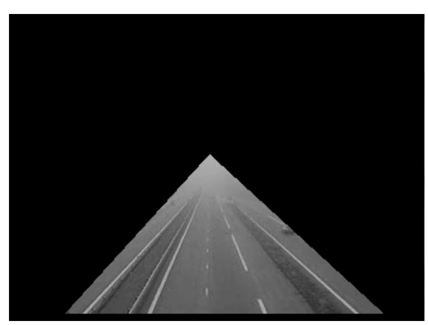

Figure 2. The region of interest, where we must calculate the contrast.

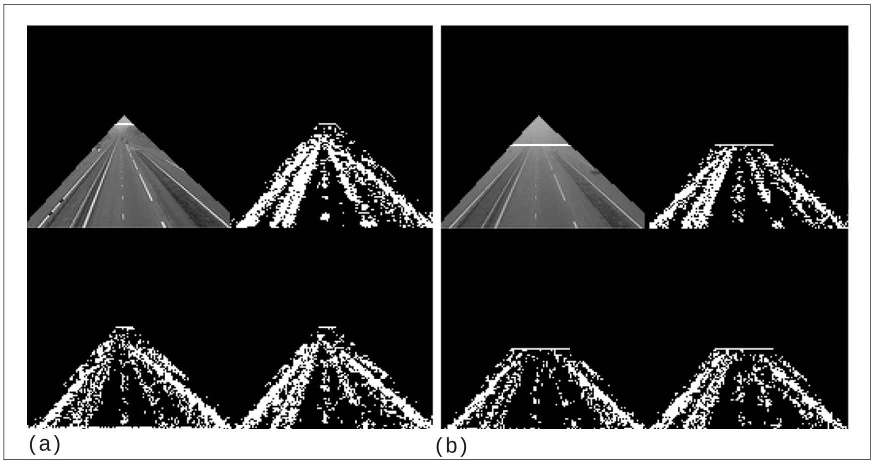

Figure 3. Results with B-spline wavelets proposed by Unser for (a) the image with a clear atmosphere and (b) the image in fog conditions.

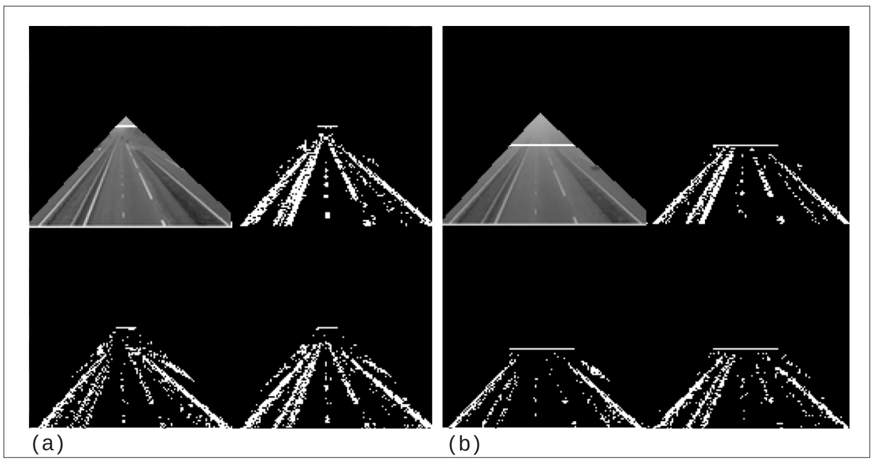

Figure 4. Results with Mallet spline for (a) the image with a clear atmosphere and (b) the image in fog conditions.

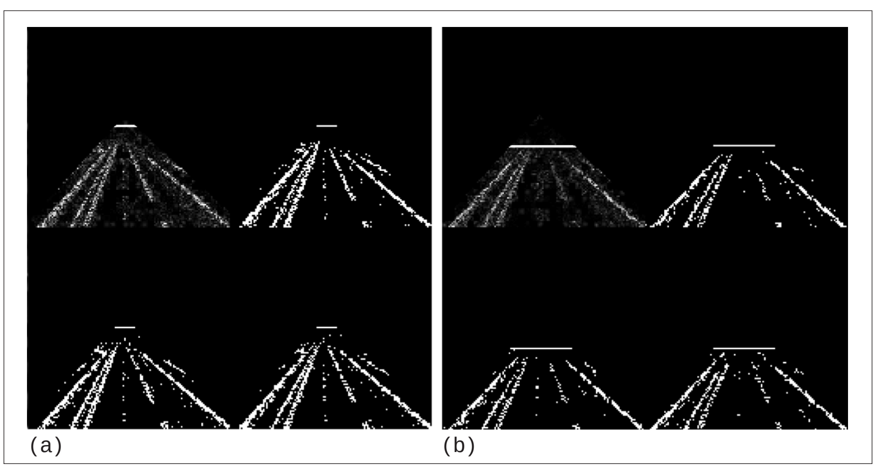

Figure 5. Results with the splines we calculated for (a) the image with a clear atmosphere and (b) the image in fog conditions.

scene, which is exactly what happens in the human visual system.

Thanks to improved information through VMS, we can reduce the number of accidents. Another possible application is an individual driving support system that improves the driver's view of the traffic scene. In this case, we could display an image showing a much brighter and sharper view of the road directly on the car's windshield or on a separate monitor integrated in the dashboard. We can achieve this by combining wavelet-based edge detection—which elaborates the significant edges of an image influenced by fog—and an additional contrast enhancement based on Jian Lu's method,[11] including an inverse transform of the modified coefficients.

References

1. J.H. Hogema et al., "Evaluation of an Automatic Fog-Warning System," *Traffic Eng. and Control*, 1996, pp. 629–632.
2. L.S. Stone and P. Thomson, *Effect of Contrast on Human Speed Perception*, Tech. Report 103898, NASA, Calif., Dec. 1992.
3. I.D. Brown, "Driving in Fog," *Symp. Eyes on the Road*, England, Nov. 1983, pp. 26–27.
4. E.A. DeYoe and D.D. Van Essen, "Concurrent Processing Streams in Monkey Visual Cortex," *Trends in Neuro Sciences*, Vol. 11, No. 5, 1988, pp. 219–226.
5. S. Mallat and W.L. Hwang, "Singularity Detection and Processing with Wavelets," *IEEE Trans. Information Theory*, Vol. 38, No. 2, March 1992, pp. 617–643.
6. D. Marr and E. Hildreth, "Theory of Edge Detection," *Proc. Roy. Soc. London*, Vol. B207, 1980, pp. 187–217.
7. R.A. Young, "The Gaussian Derivative Model for Machine Vision: Visual Cortex Simulation," *J. Opt. American Soc.*, July 1987.
8. V. Torre and T.A. Poggio, "On Edge Detection," *IEEE Trans. PAMI*, Vol. 8, 1986, pp. 147–163.
9. M. Unser, A. Aldroubi, and M. Eden, "On Asymptotic Convergence of B-Spline Wavelets to Gabor Functions," *IEEE Trans. Information Theory*, Vol. 38, No. 2, March 1992, pp. 864–872.
10. S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 14, No. 7, July 1992, pp. 710–732.
11. J. Lu, D.M. Healy Jr., and J.B. Weaver, "Contrast Enhancement via Multiscale Gradient Transformation," *Proc. Int'l Conf. Image Processing*, Sept. 1996.

Theory of B-splines

The central continuous B-spline of order n is denoted $b^n(x)$ and can be generated by repeated n convolutions of a B-spline of order 0:

$$b^n(x) = b^0 * b^{n-1} \quad (1)$$

where the 0th order B-spline $b^0(x)$ is the window function with support [-1/2, 1/2]. The spectrum of this function is given by

$$\hat{\beta}^n(\omega) = \left[\hat{\beta}^0(\omega)\right]^{n+1} = \left(\frac{\sin\left(\frac{\omega}{2}\right)}{\frac{\omega}{2}}\right)^{n+1} = \operatorname{sinc}^{n+1}\left(\frac{\omega}{2}\right) \quad (2)$$

For an integer $m \geq 1$, $b^n_m(x)$ is the nth-order B-spline dilated by scale factor m:

$$\beta^n_m(x) = \frac{1}{m}\beta^n\left(\frac{x}{m}\right) \quad (3)$$

Consider now a sequence of embedded polynomial spline function $S^n_{(i)}$, $i \in Z$ (n, an odd integer, is the degree of the polynomial). For every $i \in Z$, $S^n_{(i)}$ is a subspace of L^2 and belongs to class C^{n-1}. On each interval $[k2^i, (k+1)2^i]$ ($k \in Z$), $S^n_{(i)}$ are equivalent to a polynomial of order n:

$$S^n_{(i)} = \left\{ g^n_{(i)}(x) = \sum_{-\infty}^{+\infty} c_i(k)\beta^n_{2^i}(x - 2^i k), (x \in \Re, c_i \in 1_2) \right\} \quad (4)$$

$S^n_{(i)}$, $i \in Z$ constitutes a multiscale or multiresolution approximation of L^2, for example:

$$S^n_{(i+1)} \subset S^n_{(i)}, i \in Z \quad (5)$$

and

$$\lim_{i \to \infty} S^n_{(i)} = L^2 \quad (6)$$

Any signal $f \in L^2$ can be represented as a weighted sum of translated and dilated B-splines and are completely determined by coefficients $c_i(k)$.

The nesting relation (Equation 5) for $i = 1$ can be written $S^n_1 \subset S^n_0$. Because $\beta^n(x/2) \in S^n_0$, we have the following two-scale relation:

$$\frac{1}{2}\beta^n\left(\frac{x}{2}\right) = \sum_{-\infty}^{+\infty} h_k \beta^n(x - k) \quad (7)$$

or its equivalent in the Fourier domain :

$$\hat{\beta}^n(2\omega) = H(e^{i\omega})\hat{\beta}^n(\omega) \quad (8)$$

where

$$H(e^{i\omega}) = \sum_k h_k e^{ik\omega} = \frac{\hat{\beta}^n(2\omega)}{\hat{\beta}^n(\omega)} = \frac{\operatorname{sinc}^{n+1}\omega}{\operatorname{sinc}^{n+1}\frac{\omega}{2}} =$$

$$\left(\cos\frac{\omega}{2}\right)^{n+1} = \left(\frac{e^{i\frac{\omega}{2}} + e^{-i\frac{\omega}{2}}}{2}\right)^{n+1} \quad (9)$$

$$H(e^{i\omega}) = \sum_{j=0}^{n+1} \frac{1}{2^{n+1}} \binom{n+1}{j} e^{i\left[j - \frac{(n+1)}{2}\right]\omega} \quad (10)$$

The FIR filter coefficients are the coefficients of $H(e^{i\omega})$ in the base $(e^{ik\omega})$, so that if Equation 10 is projected on $e^{ik\omega}$, we directly get the FIR coefficients h_k for n odd:

$$h_k = \begin{cases} \frac{1}{2^{n+1}} \binom{n+1}{\frac{n+1}{2} + k}, & if |k| \leq n+1 \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

For n even, if the central B-spline $b^n(x)$ is shifted by half a sampling step with respect to the origin; it is then called a causal B-spline. For simplicity, we still denote this function by $b^n(x)$ with the abuse of notation. In this case, we have:

$$H(e^{i\omega}) = \sum_k h_k e^{ik\omega} = e^{i\frac{\omega}{2}}\left(\cos\frac{\omega}{2}\right)^{n+1} = e^{i\frac{\omega}{2}}\left(\frac{e^{i\frac{\omega}{2}} + e^{-i\frac{\omega}{2}}}{2}\right)^{n+1} \quad (12)$$

$$H(e^{i\omega}) = \sum_{j=0}^{n+1} \frac{1}{2^{n+1}} \binom{n+1}{j} e^{i\left[j - \frac{n}{2}\right]\omega} \quad (13)$$

In this case, we can also obtain the the FIR filter coefficients through projection of Equation 13 on the vectors of the base $(e^{ik\omega})$, so that for n, even the corresponding FIR filter coefficients are

$$h_k = \begin{cases} \frac{1}{2^{n+1}} \binom{n+1}{\frac{n}{2} + k}, & \text{if } -\frac{n}{2} \leq k \leq \frac{n}{2} + 1 \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

The two-scale relation (Equation 7) makes B-spline very attractive for wavelet theory and multiscale analysis. These properties can be extended to the discrete space so that the discrete sampled B-spline can define a family of multiresolution approximation spaces of 1^2:

$$S^n_i = \left\{ v \in 1^2, v(j) = \sum_{k \in Z} c_i(k) b^n_i(j - 2^i k), c_k \in 1^2 \right\} \quad (15)$$

where $b^n_i(j) = \beta^n(j/2^i)$ represents the discrete sampled spline with n odd and $j \in Z$. 1^2 consists of all summable discrete sequences with finite energy.

The embedding or refinement relations between the discrete spaces s^n_i still holds the same properties for the continuous polynomial spline of order n:

$$\lim s^n_i = 1^2, s^n_{i+1} \subset s^n_i, (\forall i \in Z) \quad (16)$$

Christoph Busch is head of the department of Security Technology for Graphics and Communication Systems at the Fraunhofer Institute for Computer Graphics in Darmstadt, Germany. He has worked in the fields of image and video analysis, traffic analysis, telecommunication, and telemedicine. He now works on projects dealing with copyright protection, digital watermarking, and conditional access for interactive multimedia services on the Internet and broadcasting services. He received his PhD in computer graphics from the Technical University Darmstadt. Contact him at the Computer Graphics Center, Rundeturmstr. 6, D-64283 Darmstadt, Germany; busch@igd.fhg.de.

Eric Debes is a PhD student in the Signal Processing Laboratory at the Swiss Federal Institute of Technology, Lausanne. His research interests include image and video coding, real-time image and video processing on general-purpose processors, instruction-level parallelism, and new multimedia architectures. He received an MS in electrical engineering and computer science from Supélec, France, and an MS in electrical engineering from the Technical University Darmstadt, Germany. Contact him at the Signal Processing Laboratory, Swiss Federal Inst. of Tech., CH-1015 Lausanne, Switzerland; eric.debesdebes@epfl.ch.