

Digital Signal 압축 기술의 섬유 산업에의 응용

김 주 용

1. 서 론

디지털기술의 발달과 더불어 빠르게 진행되고 있는 정보화의 물결은 오랜 기간동안 유지되어 온 산업의 기본구조를 근본적으로 바꿔놓고 있다. 이런 변화는 비단 컴퓨터·전기·전자공학에 의해 눈부신 발전을 거듭하고 있는 정보통신 산업뿐만 아니라, 화학·기계·섬유산업 등 소위 전통산업 등에도 커다란 영향을 미치고 있다. 특히 오랫동안 사양으로 인식되어 왔던 섬유산업은 상대적으로 낙후된 현재의 기술 수준을 제고하기 위해 재료 / 가공 / 공정 분야 전반에 걸쳐 생산 기반의 디지털화에 총력을 기울이고 있다. 최근 경쟁적으로 개발되고 있는 섬유 기계들의 디지털화, 자동화 추세는 이런 경향을 잘 반영하고 있다[1]. 새로운 생산 기계들은 기존의 시스템을 완전히 대체하거나 또는 전체 공정의 일부분에 국부적으로 편입되는 양상을 보이고 있는데, 전자의 경우에는 새롭게 발생하는 품질문제의 효율적인 관리, 후자에는 기존의 생산시스템과의 효과적인 조화라는 새로운 문제를 야기한다. 고속화, 자동화 생산 시스템하에서 발생하는 품질 문제의 원인 색출 및 교정대책의 수립은 매우 복잡하며, 그에 수반되는 효율적인 공정 제어 / 관리 시스템도 과거와는 완전히 다른 새로운 형태이어야 한다.

지난 수년동안 시도되었던 섬유공정의 실시간 계측 / 관리 시스템의 개발 노력은 다음과 같은 문제점들로 인해 가시적인 성과를 거두지 못하고 있는 실정이다. 먼저 데이터로부터 정보를 추출

하는 방법론의 부재를 들 수 있다. 공정 관리 시스템에 의해 실시간으로 수집된 데이터에 감춰져 있는 정보의 본질과 내용은 제품의 성능과 관련되어 해석되어야 하며, 시간, 위치 관련된 변수들은 적절한 형태로 변환되어 정보의 형태로 유지되어야 한다. 다음으로는 고속화되어 가는 생산 시스템에 필수적으로 수반되는 대량의 데이터분석 알고리즘의 부재이다. 전통적인 오프라인 분석법내지는 대량의 데이터를 몇 개의 통계량으로 변환한 후 폐기하는 방법은 새로운 생산시스템의 관리시스템의 소프트웨어로서는 적합하지 못하다. 따라서 새로운 생산시스템에 부합하는 효율적인 공정관리시스템의 수립은 필수불가결하며 시급히 확립되어야 하는 디지털 공정시대의 핵심 과제이기도 하다. 이에 본 논문에서는 공정의 디지털화에 필연적으로 수반되는 공정정보 시스템의 구축 원리와 그 구체적 방법론으로서의 데이터 압축 기술에 대해 논해보기로 한다.

2. 데이터 압축 기술의 원리와 종류

압축의 사전적 정의는 불필요한 정보를 삭제하고 반복되는 정보를 간략히 기술하여 같은 양의 정보를 더 작은 크기의 데이터로 표현하는 기술이라 할 수 있다[2]. 협의의 압축기술은 전술한 사전적 정의를 실현하는데 필수적인 방법론 내지는 알고리즘을 지칭하나, 넓은 의미의 압축기술은 데이터의 특성을 파악하여 유용한 정보의 형태로 변환하는 제반 기술을 의미한다. 본 논문의 연구 대상인 광의의 압축기술은 데이터

의 특징 분석, 모델링, 통계적 처리 및 모사 등 여러 기술들을 통칭하는 개념으로서 현재 전기·전자공학 분야에서 활발히 진행되고 있는 데이터 전송 및 표준화와 관련된 일련의 연구보다는 좀더 포괄적인 개념이라고 할 수 있다. 이러한 압축 기술은 현존하는 시스템 저장 용량의 물리적 한계 극복, 원격 관리 / 조정 시스템의 구현을 위한 데이터 전송시간의 최소화, 그리고 효율적인 정보 추출 및 관리라는 당면한 과제를 동시에 해결할 수 있는 유일한 대안이다. 데이터 압축 기술의 핵심은 유용한 정보의 손실 없이 데이터의 크기를 최소화하는데 달려 있다. 압축 시스템의 성능을 평가하는 기준으로는 압축률과 압축전후의 데이터 특성 차이 등을 들 수 있으나, 압축률은 출력 데이터에 대한 입력데이터의 비로써 압축전후의 데이터 품질에 대한 정보는 결여되어 있으므로, 최소 요구 수준에서 평가된 압축률을 압축 시스템 성능 평가의 기준으로 삼는 것이 일반적이다.

데이터 전송이나 압축기술은 크게 무손실 기법과 손실 기법의 2가지 종류로 구별될 수 있는데, 전자는 압축한 데이터를 복원했을 때 복원한 데이터가 압축전의 데이터와 완전히 일치하는 것을 말한다. 무손실 기법은 압축률의 극대화보다 데이터의 정확한 재현이 더 중요할 때 사용되는 방법으로 의학용 영상 등이 이에 속한다고 볼 수 있다. 후자는 복원한 데이터가 압축전의 데이터와 일치하지 않는 기법을 말한다. 이 기법은 대체로 연속 매체- 음향, 비디오, 동영상을 압축하는데 적당하다. 물론 압축 / 복원 후의 정보가 본래의 정보와 확연히 드러날 정도의 차이를 보이지는 않으며, 약간의 정보 손실을 감수하면서 높은 압축률을 얻는 방법이다. 일반적으로 제조 공정에서 요구되는 형태의 압축 기술은 주로 후자이며, 압축하고자 하는 대상 정보의 고유한 특성을 이용하므로 무손실 기법보다 더 좋은 압축률을 달성한다. 물론 손실 기법과 무손실 기법 사이의 관계는 상호 배타적인 것이 아니며, 최고의 압축률을 얻기 위해서는 손실 기법과 무손실 기법을 적절히 조합하여

사용할 수 있고, 또 정해진 알고리즘에 의해 순차적으로 또는 계층적으로 2가지 기법을 복합적으로 적용할 수 있다.

데이터 압축의 과정은 압축 알고리즘의 수립으로부터 시작되는데, 무손실 기법의 경우에는 데이터의 완벽한 복원이 주된 목적이므로 어떤 것이 유용한 정보인가에 관한 결정 기준은 크게 중요하지 않게 되는 반면, 손실기법의 경우에는 압축 대상 데이터에 내재되어 있는 유용한 정보와 불필요한 정보간의 분리가 매우 중요한 과제로 대두되게 된다. 특히 공정정보 시스템에서와 같이 데이터가 제품의 최종 속성과 관련되어 해석되어야만 하는 상황에서서의 명확한 의사 결정 기준의 수립은 시스템의 성능을 결정적으로 가늠하는 중요한 척도가 될 것이다.

현재까지 알려진 압축 기법들을 살펴보면, 먼저 허프만 부호화(Huffmann coding)로서 통계적인 부호화 방식을 채택하고 있다. 즉, 빈번히 발생하는 데이터의 코드는 적은 수의 비트로 표현하고, 빈번하지 않은 데이터는 상대적으로 많은 비트수로 표현하여 전체 데이터의 크기를 줄이는 방식으로 압축순서는 압축하고자 하는 데이터에서 발생회수를 검사하여 발생회수를 발생회수표에 저장한 후 이것을 기반으로 각 데이터에 최적화된 코드를 부여하는 방식으로 정화상, 동영상의 압축에 주로 사용되고 있다. 두번째로는 반복 길이(RL: run-length coding) 기법으로 반복되는 데이터의 블록을 하나의 값과 그 값의 반복회수로 나타내는 것으로써, 예를 들면 아래와 같이 '0'이 7번 반복되는 수열을 '0'과 반복회수를 나타낼 수 있는 플래그(n)와 반복횟수 '7'로 표현하는 방식이다.

압축할 데이터 5 7 0 0 0 0 0 0 0

압축 후 데이터 5 7 0 n 8

위와 같은 방법을 사용하면 10바이트 크기의 데이터를 5바이트 크기의 데이터로 압축 가능하며, 반복되는 회수가 클 때 큰 효과를 보인다. 세번째는 보간 압축(interpolative compression) 기법으로서 특정한 보간법을 이용하여 데

이터와 데이터간에 생략된 데이터를 채워넣는 방법이다. 데이터간의 차이가 많지 않을 때 효과적인 방법이다. 네번째로는 신호를 배경과 물체의 2가지 영역으로 분리한 후, 비교적 천천히 변하는 배경 신호를 적당한 모델로써 정형화하고 시간에 따라 빠르게 변하는 물체의 표현에만 집중하는 방식이다. 이 방법은 비교적 최근에 많이 시도된 방식으로써 시간영역에서 구축되는 시계열모델(선형필터), 최적 필터모델(확률 필터), 프랙탈모델(비선형모델)과 주파수 영역에서 구축되는 파워스펙트럼 매칭모델, 웨이블릿 모델 등 여러 가지 형태가 보고되고 있다[3]. 이 방법은 배경 신호를 적절하게 표현할 수 있는 모델의 수립이 최대의 관건이며, 신호의 속성이 매우 빠르게 변하는 경우에는 적절한 모델의 수립이 매우 어렵게 되므로, 복잡한 수학적 기술을 동원해야 하는 경우가 많다. 이외에도 새로 부호화할 정보가 이전정보와 강한 상관관계가 성립할 때 예측을 통해 기록하는 방법, 또는 차이만을 기록하는 방법, 시간이나 주파수 영역에서 기저함수를 이용하는 방법 등 많은 형태의 방법들이 있다.

최근에 가장 많은 주목을 끌고 있는 압축 방법 중의 하나가 웨이블릿 변환을 이용한 방법인데, 원래의 신호를 시간-주파수 영역에서 효율적으로 표현할 수 있는 기저함수인 웨이블릿을 이용하여 신호를 분해한 후 상대적으로 낮은 값을 나타내는 웨이블릿들은 제거함으로써 고도의 압축률을 달성하는 방식이다. 다른 압축법에 비해 그 효과가 탁월하고 또한 신호를 여러 주파수로 분해하거나 분리할 수 있는 장점 때문에 많은 주목을 받고 있다. 웨이블릿을 응용한 압축기술은 비교적 새롭게 성장하고 있는 분야이며, 섬유 공학에 적용된 예는 많지 않다. 1996년에 Jasper 등은 웨이블릿변환을 이용하여 직물의 결점을 검출하려는 시도를 하였다[4]. 이 연구에서 사용된 적응 웨이블릿은 결점 부분의 신호에만 예민하게 반응하도록 설계되었는데, 이는 앞에서 언급된 최적 베이스시스 설계의 한 예라 할 수 있다. 이후

Jeong 등은 직물 이미지의 압축에 관한 연구에서 웨이블릿 변환 기술에 의해 원래의 이미지가 커다란 외관의 변화없이 90% 이상 압축 가능함을 보고하였다[5].

3. 웨이블릿 변환

앞에서 전술했던 여러 압축 기술중 가장 최근에 주목을 끌고 있는 신기술인 웨이블릿 변환을 이용한 압축기술에 대해 알아보기로 한다.

3.1. 최적 베이스시스의 설계 기술

이 절에서는 웨이블릿 행렬(wavelet matrix)을 분해하는 원리와 방법에 대해 서술한다[6]. 일반적인 m 대역 이산 웨이블릿 변환의 경우, 웨이블릿 필터계수는 1개의 저주파 대역 필터와 $m-1$ 개의 고주파 대역필터로 구성된다. 실제 시스템에서 흔히 발견되는 양대역 웨이블릿필터는 이의 특수한 경우이다. 만약 3대역 이산 웨이블릿의 경우를 살펴보면, 1개의 저대역 필터(h)와 2개의 고대역 필터(g)로 구성되며, 각각,

$$\begin{aligned} h &= (h_0, \dots, h_5), \mathbf{g}^{(1)} = (\mathbf{g}_0^{(1)}, \dots, \mathbf{g}_5^{(1)}), \mathbf{g}^{(2)} \\ &= (\mathbf{g}_0^{(2)}, \dots, \mathbf{g}_5^{(2)}) \end{aligned}$$

으로 구성되는 필터계수를 갖게 된다($n=6$). 원래의 신호 $\mathbf{x}_{(0)}$ (최고 주파수대역)는 다음 계산에 의해 다음 주파수대역신호 $\mathbf{x}_{(1)}$ 으로 변환된다.

$$\begin{aligned} \mathbf{x}_{(1)} &= \mathbf{W}^{(0)} \cdot \mathbf{x}_{(0)} = \mathbf{P}^{(0)} \cdot \mathbf{W}_{cir}^{(0)} \cdot \mathbf{x}_{(0)} \\ &= (a_{10} a_{11} a_{12} d_{10}^{(1)} d_{11}^{(1)} d_{12}^{(1)} d_{10}^{(2)} d_{11}^{(2)} d_{12}^{(2)})^T \end{aligned} \quad (1)$$

$$\text{여기서, } a_{j+l,k} = \sum_{i=0}^{n-1} h_i a_{j,mk+i}$$

$$d_{j+l,k} = \sum_{i=0}^{n-1} g_i^{(b)} a_{j,mk+i} (b=1, \dots, m-1)$$

순열행렬(permutation matrix)과 원 행렬(circulant matrix)의 곱은 다음처럼 표시되며,

$$P^{[0]}W_{cir}^{[0]} = \quad (2)$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times$$

$$A_0 = \begin{bmatrix} h_0 & h_1 & h_2 \\ g_0^{(1)} & g_1^{(1)} & g_2^{(1)} \\ g_0^{(2)} & g_1^{(2)} & g_2^{(2)} \end{bmatrix} \quad (5)$$

$$A_1 = \begin{bmatrix} h_3 & h_4 & h_5 \\ g_3^{(1)} & g_4^{(1)} & g_5^{(1)} \\ g_3^{(2)} & g_4^{(2)} & g_5^{(2)} \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} h_0 & h_1 & h_2 & h_3 & h_4 & h_5 & 0 & 0 & 0 \\ g_0^{(1)} & g_1^{(1)} & g_2^{(1)} & g_3^{(1)} & g_4^{(1)} & g_5^{(1)} & 0 & 0 & 0 \\ g_0^{(2)} & g_1^{(2)} & g_2^{(2)} & g_3^{(2)} & g_4^{(2)} & g_5^{(2)} & 0 & 0 & 0 \\ 0 & 0 & 0 & h_0 & h_1 & h_2 & h_3 & h_4 & h_5 \\ 0 & 0 & 0 & g_0^{(1)} & g_1^{(1)} & g_2^{(1)} & g_3^{(1)} & g_4^{(1)} & g_5^{(1)} \\ 0 & 0 & 0 & g_0^{(2)} & g_1^{(2)} & g_2^{(2)} & g_3^{(2)} & g_4^{(2)} & g_5^{(2)} \\ h_3 & h_4 & h_5 & 0 & 0 & 0 & h_0 & h_1 & h_2 \\ g_3^{(1)} & g_4^{(1)} & g_5^{(1)} & 0 & 0 & 0 & g_0^{(1)} & g_1^{(1)} & g_2^{(1)} \\ g_3^{(2)} & g_4^{(2)} & g_5^{(2)} & 0 & 0 & 0 & g_0^{(2)} & g_1^{(2)} & g_2^{(2)} \end{bmatrix}$$

위에서 언급된 원 행렬 W_{cir} 은 다음과 같이 표시할 수 있다.

$$W_{cir} = \begin{bmatrix} A_0 A_1 & 0 \\ 0 & A_0 A_1 \\ A_1 & 0 A_0 \end{bmatrix} \quad (7)$$

웨이블렛 베이스함수와 MRA가 존재하기 위한 필터계수의 조건은 다음과 같다.

- 1) 직교성 $\sum A_k A_{k+i}^T = \delta_{0i} I$, 여기서 δ_{0i} 는 크로네커델타이고 I 는 단위 행렬
- 2) 정규조건(regularity condition) $\sum_k h_k = \sqrt{3}$,
- 3) 로튼행렬(Lawton Matrix)

$$M_{ij} = \sum_k h_k h_{k+j-mi}$$

여기서, 행렬의 고유치는 반드시 1이어야 한다.

위의 3가지 조건을 만족하는 행렬은 어떤 것이든 웨이블렛 행렬이 될 수 있으며, 계수의 개수나 값을 변화시키면서 특정한 성질을 갖는 필터를 설계할 수 있다. 1994년에 Naoki는 여러 종류의 신호를 분리할 수 있는 성능을 가진 웨이블렛 필터를 설계하였으며, 이를 국부 변별 베이스(local discriminant basis)라 지칭하였다[7]. 데이터 압축과 같은 특정한 용도에 적합한 웨이블렛 베이스(wavelet basis)도 다음과 같은 최적화 알고리즘을 통해 쉽게 만들어질 수 있다.

주어진 임의의 정수 $m \geq 2$ 와 $k \geq 1$ 이고, 직교조건을 만족시키는 임의의 $m \times mk$ 행렬 A 는

$$A = (A_0 A_1 \dots A_{k-1}) = H_o F_1 o \dots o F_q \quad (8)$$

여기서 기호 o 는 다음과 같이 표현되는 “다항

원 행렬, W_{cir} 은 첫째 행은 저주파 대역필터로 나머지 2행은 고주파 대역필터로 이루어지는 3×6 의 행렬이며, 다음과 같이 나타낼 수 있다.

$$W_{cir} = \begin{bmatrix} A_0 A_1 \dots A_q & 0 & \dots & 0 & 0 \\ 0 & A_0 \dots A_{q-1} A_q & \dots & 0 & 0 \\ & & \vdots & & \\ A_2 A_3 \dots & 0 & 0 & \dots & A_0 A_1 \\ A_1 A_2 \dots & 0 & 0 & \dots & 0 A_0 \end{bmatrix} \quad (3)$$

위의 행렬 A 는 다시 $(A_0 A_1 \dots A_k)$ 와 같은 여러 개의 3×3 차행렬(sub-matrix)로 분해될 수 있으며, 여기서 k 는 $(n/m)-1$ 이다.

따라서, 웨이블렛 행렬 A 는

$$A = \begin{bmatrix} h_0 & h_1 & h_2 & h_3 & h_4 & h_5 \\ g_0^{(1)} & g_1^{(1)} & g_2^{(1)} & g_3^{(1)} & g_4^{(1)} & g_5^{(1)} \\ g_0^{(2)} & g_1^{(2)} & g_2^{(2)} & g_3^{(2)} & g_4^{(2)} & g_5^{(2)} \end{bmatrix} \quad (4)$$

다음 2개의 행렬 $(A_0 A_1)$ 로 분해될 수 있다.

곱(polynomial product)을 나타낸다.

$$(B_0 B_1 \dots B_{p-1}) \circ (C_0 C_1 \dots C_{r-1}) \\ = (D_0 D_1 \dots D_{p+r-2})$$

여기서,

$$D_j = \sum_k B_k C_{j-k} \quad (9)$$

인수형 $F_j = (P_j \mathbf{I} - P_j)$, 여기서 P_j 는 대칭형 투영체 직교행렬이다($HH^T = \mathbf{I}$). 4의 필터길이를 가진 2대역 변환의 경우 직교조건이 만족되면 \mathbf{A} 는 다음과 같이 표시될 수 있다.

$$\mathbf{A} = \mathbf{H} \circ \mathbf{F}_1 = \mathbf{H} \circ (\mathbf{R}_1 \mathbf{I} - \mathbf{R}_1) \\ = [\mathbf{H} \mathbf{R}_1 \mathbf{H} (\mathbf{C} \mathbf{I} - \mathbf{R}_1)] = [\mathbf{A}_0 \mathbf{A}_1] \quad (10)$$

다음 단계는 \mathbf{A} 의 인수분해를 통해 행렬 \mathbf{H} 와 투사행렬 \mathbf{R}_i (for $i=1, \dots, k$)을 찾아내는 것이다. 전술한 정규조건에 의해 행렬 \mathbf{H} 의 첫번째 행을 모두 $1/\sqrt{m} \mathbf{1}$ (여기서 $\mathbf{1}$ 은 단위벡터)로 고정할 수 있으며 나머지 $m-1$ 행만이 직교성을 만족시키는 조건 하에서 최적화된다. 위의 직교성을 만족시키기 위해서는 나머지 $m-1$ 행이 $(\mathbf{I} - 2\mathbf{u}\mathbf{u}^T)\mathbf{T}^\nabla$ 여기에서 \mathbf{u} 는 정규화된 벡터)에 의해서 계산되어야 하며, \mathbf{T}^∇ 는 대각선 원소가 $T_{ii} = i-m$ 이고 나머지는 모두 1인 정규화된 위 삼각형 행렬 (upper triangle matrix)이다. 결과적으로 $m \times m$ 의 직교행렬 \mathbf{H} 는 다음과 같은 분할된 형태로 나타낼 수 있다.

$$\mathbf{H} = \begin{pmatrix} 1/\sqrt{m} \mathbf{1} \\ (\mathbf{I} - 2\mathbf{u}\mathbf{u}^T) \cdot \mathbf{T}^\nabla \end{pmatrix} \quad (11)$$

랭크 τ 의 대칭 투영행렬 (symmetric projection matrices)은 $\mathbf{R} = \mathbf{V}\mathbf{V}^T$ 로 표시될 수 있으며, 여기서 $\mathbf{V}_{m \times \tau}$ 는 정규직교성의 열 벡터를 가진 행렬이다. 웨이블릿 행렬이 nonredundant 하기 위해선 투영행렬의 랭크가 단조감소, 즉 $\text{rank}(\mathbf{R}_1) \leq \dots \leq \text{rank}(\mathbf{R}_k)$ 와 같은 형태여야하며, 이외에 랭크가 1인 일반적인 제한조건이 부가되어 다음과 같은 형태를 얻는다.

$$\mathbf{R}_i = \mathbf{v}_i \mathbf{v}_i^T \quad (12)$$

섬유기술과 산업, 제 5 권 제 1/2 호, 2001년

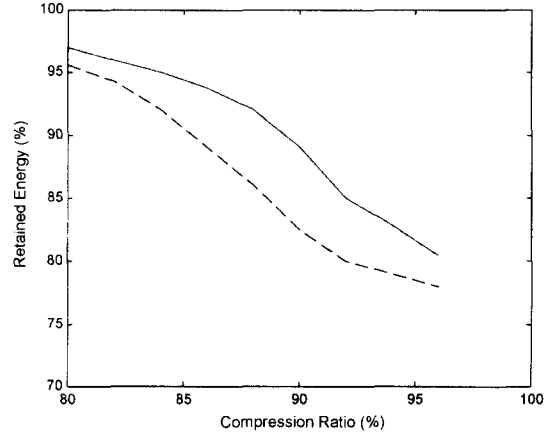


Figure 1. Comparison of the compression performances of two bases.

여기서 $\mathbf{v}_i \mathbf{v}_i^T = 1$

웨이블릿 행렬을 구성하기 위해서는 위에서 얻어진 직교행렬 \mathbf{H} 와 투영행렬들이 필요한데, 구체적으로 균일 분포 함수로부터 무작위로 만들어진 정규 벡터들 $\mathbf{v}_1, \dots, \mathbf{v}_k$, \mathbf{u} 에 의해 만들어진다. Figure 1은 전술한 알고리즘을 기초로 오픈-엔드 방식의 신호에 대해 압축률을 극대화시키도록 최적 설계된 웨이블릿(실선)의 성능을 일반 웨이블릿(점선, Daubechies 5 웨이블릿)과 비교한 것이다. 최적설계 웨이블릿이 약 80%의 압축비에 달할 때까지 월등한 압축비를 나타낼 수 있다.

3.2. 웨이블릿 패킷변환

웨이블릿 패킷변환은 웨이블릿 변환의 일반적인 형태이다. 웨이블릿 변환은 신호안에 존재하는 중요한 시간-주파수 요소들을 찾아내는 효과적인 방법이지만, 저주파 필터에 의해 여과된 신호만을 대상으로 필터를 적용하기 때문에 주파수 대역이 제한을 받게 된다. 이런 단점을 개선하기 위해 고주파영역에 있는 신호들을 더 세분화시켜 주파수 분해능을 향상시킨 방법이 바로 웨이블릿 패킷변환이다. 웨이블릿패킷 Ψ 은 제곱유한 적분 가능한 변조 파형이며, 위치와 주파수

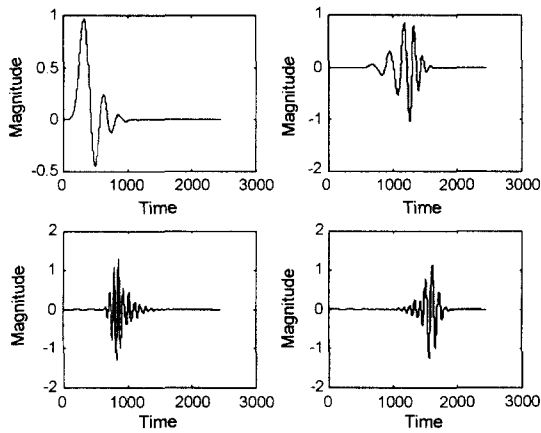


Figure 2. Wavelet packets made from a a Daubechies 10 wavelet.

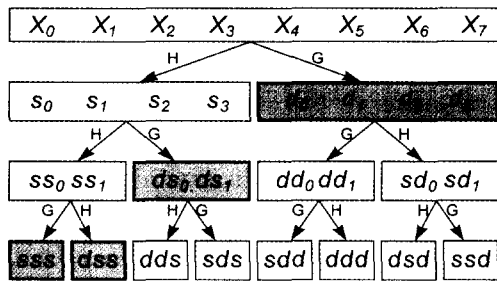


Figure 3. Binary tree structure of wavelet packet (gray box : wavelet basis).

영역 양쪽에 대해 국부화되어있다. Figure 2는 웨이블릿 패킷의 예이다[8].

웨이블릿 변환과 달리, 패킷변환은 세부 계수 (detail coefficients)들도 필터링에 의해 더 작은 주파수 대역으로 분해되므로 향상된 주파수 분해능을 보여준다. 이렇게 분해된 신호들이 주파수 대역에 따라 정렬 되면, Figure 3과 같은 나뭇가지와 같은 형태를 나타내게 된다. 그림에서 첫째 층은 변환전의 원래 신호이다. 원래의 신호는 QMF에 의해 그 크기가 1/2 씩 작아지며, 아래층의 주파수 대역도 1/2씩 줄어들게 된다. 첫째 층을 제외한 나머지 3층에 있는 계수들이 웨이블릿 패킷계수들이며, 그 중 회색 상자는 웨이블릿 변환시의 웨이블릿 계수들이다.

3.3. Best Basis 선택 알고리즘

신호가 길어짐에 따라 웨이블릿 패킷변환의 결과인 계수들의 갯수도 기하급수적으로 많아지므로, 이들의 갯수를 줄이는 여러 방법들이 필요하게 된다. 가장 효과적인 방법은 웨이블릿계수들의 값이 큰 몇 개와 작은 나머지로 구성되는 경우에 값이 큰 몇 개만을 남기고 나머지 계수들은 제거하는 경우이다. 이 경우 커다란 에너지의 손실 없이도 몇몇 계수들만을 이용해 원래의 신호를 복원할 수 있으며, 압축률도 극대화된다. 따라서, 웨이블릿 계수들을 선택하는 방법은 높은 압축률을 달성하는데 아주 중요한 요건이 된다. 위의 임무를 달성하기 위한 기준을 정보 손실 함수라 부른다. 손실함수로는 여러 종류의 함수가 있는데, 일반적으로 계산이 간편한 장점이 있는 부가 손실 함수(additive loss function)를 사용한다. 모든 유한한 길이의 신호 $f=(f_1, \dots, f_n)$ 에 대해 $M(x,y)=M(x)+M(y)$ 를 만족시키는 functional M 을 부가적인 손실 함수라 한다(여기서, $M(0)=0$). 표준 M 의 예로서 다음과 같은 엔트로피가 많이 사용된다[8].

엔트로피 : 유한한 신호 $x=\{x_j\}$ 에 대한 Shannon의 “엔트로피”는 $H(x) = -\sum p_j \log p_j$, 여기서 $p_j = |x_j|^2 / |x|^2$ 로 정의되고 부가적이지 않지만, l^2 norm $E(x) = -\sum |x_j|^2 \log |x_j|^2$ 은 부가적이다.

Coifman과 Wickerhauser 에 의해 개발된 “best basis” 선택 알고리즘은 1) 정규직교 웨이블릿의 선택과 신호와의 컨벌루션과 2) 엔트로피 기준 (norm)에 근거한 손실함수를 최소화 시키는 웨이블릿 패킷 계수들을 찾아내는 것의 두 단계로 구성된다. 이 알고리즘의 결과로 원래의 웨이블릿 패킷 계수보다 훨씬 작은 갯수의 계수가 남게 되고, 신호를 효율적으로 표현할 수 있게 된다.

4. 섬유 공업에의 응용 : 방직공정을 중심으로

지난 수 세기동안 생산설비의 고속화, 자동화 추세에 힘입어 단위시간당 생산속도는 지속적인

로 상승해왔다. 고성능 센서와 데이터 수집 시스템의 출현과 동시에 데이터 생산속도는 기하급수적으로 늘어 방적사 600 m의 길보기 굵기를 분석, 저장하는데 요구되는 파일의 크기는 무려 1.2 MB이며 보통의 생산속도에서 약 2~3분만에 생겨 나는 데이터의 양에 해당한다. 그러나, 현재의 분석 방법은 CV%와 같은 간단한 통계량이나 시공간정보를 완전히 삭제하는 스펙트로그램 등과 같이 공정에 관한 매우 제한적인 정보만을 제공하는데 머무르고 있다[9]. 현대의 고속 생산 방적 공정을 위한 효과적인 실시간 공정 관리 시스템은 대량의 데이터로부터 필수적인 핵심정보는 유지한 채로 불필요하고 반복적인 정보들만을 효과적으로 추출, 여과할 수 있어야 한다. 이러한 시스템을 위한 하드웨어의 기술은 이미 갖추어진 상태이나, 시스템의 핵심이라 할 수 있는 알고리즘의 부재로 인해 솔루션의 개발이 지연되고 있는 실정이다. 일반적으로 빠르게 변화하는 동적인 신호의 형태를 갖는 방적사의 두께 신호를 원신호 그대로의 모습으로 재현하기 위한 압축은 커다란 의미가 없으며 또 압축률을 극대화시키기도 어렵다. 즉 손실 기법의 형태로 압축하여야 하며, 한가지의 압축 기술보다는 2가지 이상의 방법을 복합적으로 적용하는 것이 더 효과적이다. 제시될 방법은 방적사의 품질을 평가하는데 전통적으로 사용되어 오던 결점 정보들을 표현하는 통계적 모델을 이용하여, 공정을 모니터링하고 결점수의 갑작스러운 증가가 발생할 경우에만 웨이블렛 변환을 시도하여 원래의 신호를 최대한 유지하는 동시에 압축률을 극대화하는 방식이다. 이때 사용되는 확률 모델은 일종의 전처리 필터 역할을 하게 되며, 불필요한 웨이블렛 변환을 방지함으로써 실시간 처리에 소요되는 시간적 낭비를 막는 역할을 하는 동시에 결점의 공간적 분포에 대한 유용한 정보도 제공한다. Figure 4는 이런 알고리즘을 도식적으로 표현한다[10].

방적사의 품질과 직접적 연관이 있는 넵(nep)과 두꺼운 부분(thick place)의 갯수, 두꺼운 부분의 길이, 두꺼운 부분의 두께 등의 결점들에 관련된 정보들은 각각 포아슨, 감마, 극한분포에

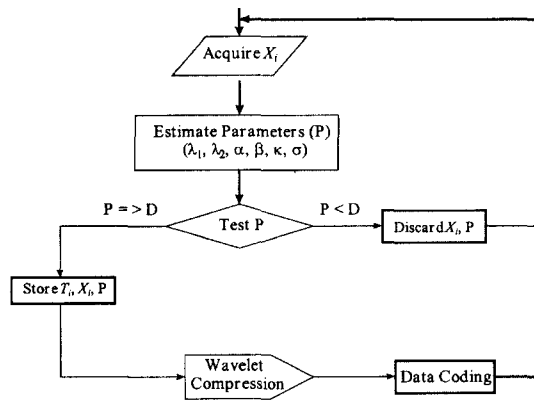


Figure 4. A flow chart for the hybrid data screening (P : estimated parameters and D : a threshold).

의해 모델링될 수 있으며, 파라미터들은 생산되는 방적사에 대해 50 m 간격으로 추정된다. 이 경우 추정되어야 하는 파라미터는 모두 6개가 되며, 미리 정해진 기준치와 비교되면서 연속적으로 관찰된다. 만약 추정된 파라미터들이 기준치를 초과하지 않았을 경우에는 그 구간(50 m)에 해당하는 모든 데이터는 저장되지 않고 제거된다. 이 경우는 불필요한 신호로 간주되어 저장되지 않는데 그 이유는 공정의 기본적인 특성 또는 반복되는 특성을 나타내는 신호에 해당되므로 새로운 정보를 담고 있지 못하고 또 확률적 모사에 의해 언제든지 재현 가능한 신호이기 때문이다. 반면, 특정 구간에서 방적사의 넵의 갯수가 기준치를 초과했을 경우, 원재료의 불균일성 혹은 공정상의 결점 때문에 이 사건이 발생했다고 간주할 수 있으므로, 사건이 발생하는 전 구간의 데이터는 즉시 저장된다. 사건의 시공간적 위치는 매우 중요한 정보이므로, 발생시간과 통계적 추정된 6개의 파라미터 그리고 웨이블렛 변환에 의해 95%로 압축된 웨이블렛 계수 등이 순서대로 파일에 저장되며 그 형태는 Figure 5와 같다. 이 파일은 공정의 역사를 핵심적으로 담고 있는 일종의 유전인자 파일이며, 결점의 분포, 결점 발생시간 그리고 결점부분의 원래 신호에 대한 정보를 모두 다 포함하고 있으면서도 그 크기는 원래 신호의 수천분의 1 미만에 불과한 압축 파일

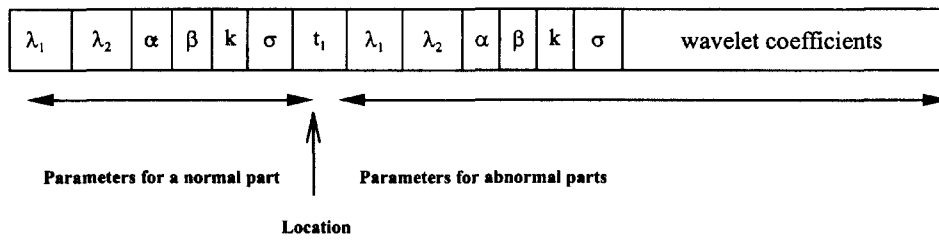


Figure 5. A coding for data storage.

Table 1. Space requirement for a continuous monitoring

Time \ Yarn	Yarn lengths (per spindle)	Number of data points (per spindle)	Required space (per spindle)
0.5 hour	4,440 m	2,220,000	8.88MB
1 hour	8,880 m	4,440,000	17.8MB
1.5 hours	13,320 m	6,660,000	26.6MB

이다.

위의 파일은 실제 방적공정에 매 스펀들마다 설치되어 생성되며, 여러 파일들을 함께 모아 데이터베이스의 형태로 관리하게 된다. 일단 데이터베이스의 형태로 저장된 위의 파일들은 그 크기가 매우 작아 중앙 컴퓨터로 전송하기 쉽고, 또 결점 발생 시간, 공간적 위치, 원래의 신호 등 거의 모든 정보를 담고 있으므로 매우 이상적인 공정 정보 시스템의 데이터베이스로 적합하다. 현재 대부분의 방적공정은 고속화되어 있으므로, 넵과 같은 결점들이 발생했을 때 실시간으로 검출하고 제거하기 위한 광학 센서들을 사용한 자동화 시스템이 구축되어 있다. 그러나, 그 시스템들은 대용량의 데이터를 처리할 수 있는 효율적인 데이터분석 알고리즘의 부재로 인해 진정한 공정 관리 시스템으로 발전하지 못하고 있는 실정이다. 연속적인 공정 관리에 요구되는 저장 공간은 Table 1에 나타나있는데, 이를 보면 빠른 속도로 발전하고 있는 데이터 수집과 저장 기술에도 불구하고 의미있는 공정 관리 시스템의 출현이 늦어지고 있는 것은 전혀 놀랄만한 일이 아님을 알 수 있다.

전술한 알고리즘을 오픈-엔드 방적 공정에 실제로 적용하여 그 성능을 평가하기 위해 10개의 방적사 패키지가 Schlaforst® 오픈-엔드 방적기

에서 생산되었으며, 그 생산속도는 분당 148 m였고 대략적인 변수는 17수였다. 각 스펀들에서 25,000 m 씩 10개의 스펀들에서 총 250,000 m에 달하는 방적사가 생산되고 Zweigle G-580® 광학용 균제도 측정기에 의해 그 두께가 매 2 mm마다 연속적으로 측정되었다. 수집된 데이터의 개수는 스펀들당 7,500,000개이고 이는 약 50 MB에 해당하는 저장공간을 요구한다. 10개의 스펀들에서는 약 500 MB에 달하므로 물리적인 저장공간의 한계에 순식간에 도달해버리게 된다. 그러나, 전술한 알고리즘을 적용하면 작은 개수의 품질 파라미터와 웨이블렛 계수로 구성된 공정정보 파일이 얻어지며, 이 실험의 경우 약 5 KB의 작은 저장공간만이 요구되고, 물리적 압축률은 약 십만분의 일에 달하였다.

5. 결 론

고속화, 대량 생산화되어 가는 섬유생산공정은 자동화된 관리 시스템에 의해 연속적으로 감시되고 효율적으로 제어되어야 한다. 좀 더 구체적으로 디지털 시대의 공정 관리 시스템은 단 시간에 생산되는 많은 양의 데이터를 효과적으로 처리하여 핵심 정보만을 추출, 여과, 압축하는 기능을 필수적으로 보유하여야 하며, 데이터

베이스의 형태로 정보화되어야 한다. 웨이블릿 변환 기술을 이용한 새로운 공정 정보 시스템은 이를 실현하는데 필요한 효율적인 방법론을 제공한다.

참고문헌

1. W. Oxehnam, ITMA Review 99: A Review, *Textile Progress*, P1-30(1999).
2. www.incheon.ac.kr
3. G. Strang and N. Truong, "Wavelets and Filter Banks", 1st Ed., Wellesley-Cambridge Press, 1996.
4. W. J. Jasper, S. J. Garnier, and H. Potlapalli, *Text. Res. J.*, **35**(11), 3140(1996).
5. S. H. Jeong and E. H. Park, *J. Korean Fiber Soc.*, **36**(6), 445(1999).
6. R. Turcajova, and J. Kautsky, *Numerical Algorithms*, **41**, 27(1994).
7. S. Naoki, Ph.D. Thesis, Dept. of Mathematics, Yale University, 1994.
8. R. Coifman, and M. V. Wickerhauser, *IEEE Trans. on Inf. Theory*, **38**(2), 713(1992).
9. R. Furter, "Evenness Testing in Yarn Production". Part1, The Textile Institute, Manchester, England, 1982.
10. J. Kim, Fiber and Polymer Sci. Program, NCSU, 1998.