C군의 네버엔딩 스토리,

디지털 영상처리의 이해 - 10

스산한 찬바람에 문득문득 어깨가 움츠러듭니다. 독자 여러분 모두께 곧 다가올 겨울 동안 감기 따위로 아픈 일 없는 무탈한 겨울나기를 빕니다. 지난 편에서는 '디지털 영상 확대' 고급편인 Bicubic 영상 확대 방법을 더욱 쉽게 이해하기 위한 발판으로 Bilinear 방식을 소개해드렸습니다. 이번 편에서는 '디지털 영상 확대'의 종결자인 Bicubic 방식을 살펴보겠습니다.

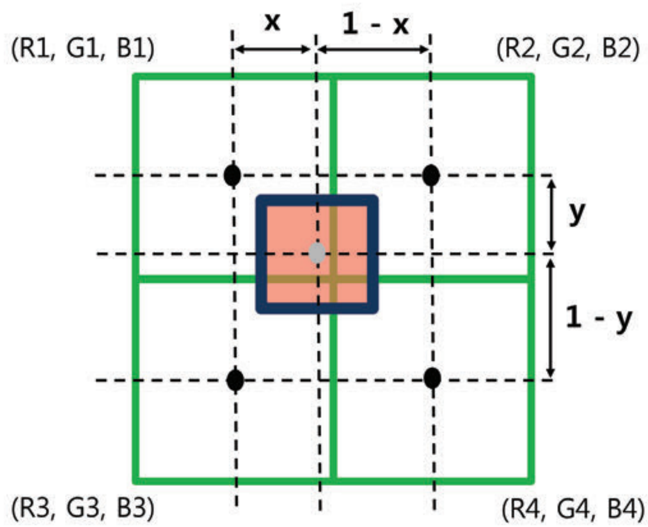

그림 1. Bilinear 방식에서 원본 영상 화소(초록색 테두리)와 확대 영상 화소(파란색 테두리)의 상대적 대응 관계

[수식 1]

$$(R', G', B') = (1-x) \cdot (1-y) \cdot (R1, G1, B1) + (1-x) \cdot y \cdot (R3, G3, B3)$$
$$+ x \cdot (1-y) \cdot (R2, G2, B2) + x \cdot y \cdot (R4, G4, B4)$$

(R', G', B') : 확대 영상 화소의 색 ([그림 1]의 파란색 테두리 화소)

[그림 1]과 [수식 1]은 지난 편에서 설명된 Bilinear 방식의 보간법을 나타내고 있습니다. Bilinear 방식을 다시 간단히 상기하면 [그림 1]의 파란 테두리 사각형은 원본 영상과 확대 영상의 크기를 맞춰 1:1 대응시켰을 때 원본 영상 화소에 대한 확대 영상 화소의 상대적 관계를 하나의 화소에 관해 표시한 것이며, [수식 1]은 [그림 1]의 확대 영상 화소의 색인 (R', G', B')를 결정하기 위해 인접한 원본 영상 화소 4개를 이용하여 (R', G', B')값을 보간하는 식입니다. 식의 구조를 보면 직감적으로 느끼시겠지만 Bilinear라는 이름이 붙은 이유가 수평/수직 방향으로 선형보간(Linear Interpolation)을 순차적으로 두 번 수행하기 때문입니다. 그렇다면 Bicubic 방식은 무엇일까요? 지난 편에서도 말씀드렸지만, Bicubic 방식도 수평/수직 방향의 순차적 보간을 두 번 수행하는 것은 Bilinear 방식과 동일합니다. 하지만 Bilinear가 일차식을 이용한 선형보간 방식인 반면 Bicubic은 3차식(Cubic이라는 단어가 암시하듯이)을 이용한 비선형 보간 방법입니다. 우선, Bicubic의 3차 보간식의 개념을 [그림 2]를 통해서 설명하도록 하겠습니다.

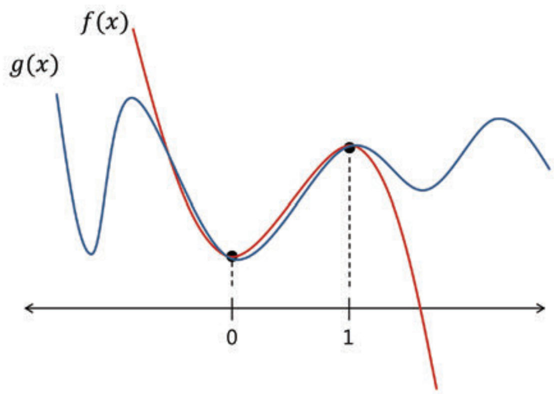

그림 2. 함수 g(x)를 $0 \le x \le 1$ 구간에서 3차식 f(x)로 근사

[그림 2]와 같이 파란 점선으로 표시된 임의의 함수 g(x)가 있고, 이 함수가 $0 \le x \le 1$ 구간에서 급격한 변화가 없는 단순 증가/감소 함수라면, 해당 구간에서 함수 g(x)의 추세에 맞춘 3차 함수 f(x)를 통해 g(x)를 근사화할 수 있습니다. g(x)를 근사화한 함수 f(x)를 구하기 위한 조건은 x = 0, 1에서 g(x)와 g(x)의 도함수 g'(x)가 갖는 값을 f(x)와 f'(x)가 그대로 갖는다고 아래와 같이 가정하는 것입니다.

$$f(0) = g(0),\ f(1) = g(1),\ f'(0) = g'(0),\ f'(1) = g'(1)$$

$0 \le x \le 1$ 구간에서 f(x)로 근사화하려는 단순 증가/감소 함수인 g(x)를 알고 있다면, g(0), g(1), g'(0), g'(1)는 이미 알고 있는 값이 되므로, 임의의 함수 g(x)를 근사화하는 3차식 f(x)를 구하는 방법은 다음과 같습니다. 우선, [수식 2]와 같이 일반적 형태의 3차 함수 f(x)를 정의합니다.

[수식 2]

$$f(x) = ax^3 + bx^2 + cx + d$$

f(x)가 정의되면 그 도함수는 미분을 통해 [수식 3]과 같이 정리됩니다.

[수식 3]

$$f'(x) = 3ax^2 + 2bx + c$$

[수식 2]와 [수식 3]에 앞에서 설정한 조건들을 대입하여 정리하면 다음과 같습니다.

[수식 4]

$$
\begin{aligned}
f(0) &= d \\
f(1) &= a + b + c + d \\
f'(0) &= c \\
f'(1) &= 3a + 2b + c
\end{aligned}
$$

앞의 다항식들을 풀어서 a, b, c, d에 대해 정리하면 [수식 5]와 같고, 이 계수들을 [수식 2]에 대입하면 $0 \le x \le 1$ 구간에서 g(x)를 근사화한 3차식 함수 f(x)가 구해집니다.

[수식 5]

$$
\begin{aligned}
a &= 2f(0) - 2f(1) + f'(0) + f'(1) \\
b &= -3f(0) + 3f(1) - 2f'(0) - f'(1) \\
c &= f'(1) \\
d &= f(0)
\end{aligned}
$$

그렇다면 앞의 3차식을 이용한 근사화와 Bicubic 방식의 영상 확대가 무슨 관계일까요? [그림 3]을 통해서 이 상관관계를 설명하겠습니다.

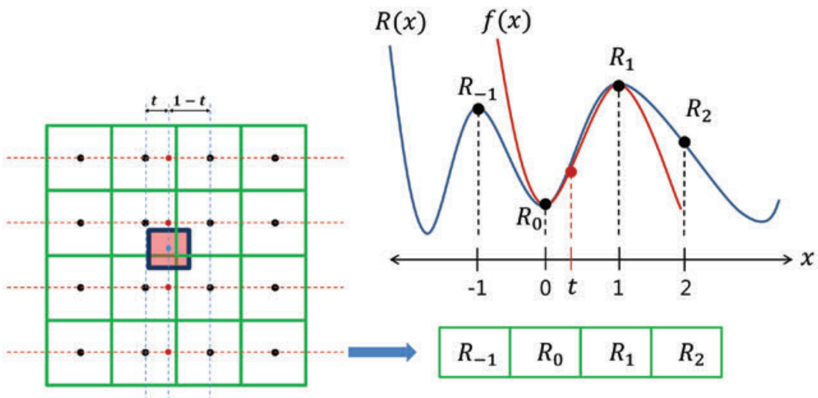

그림 3. Bicubic 방식의 3차 보간식과 화소값의 관계

[그림 3] 좌측의 초록 테두리 사각형들 내부에 있는 파란 테두리 사각형은 [그림 1]의 경우와 동일하게 원본 영상 화소에 대해 임의의 확대 영상 화소 하나를 상대적으로 나타낸 것에 해당되며, 확대 영상 화소인 파란 테두리 사각형의 색을 나타내는 *(R, G, B)*값은 원본 영상 화소인 초록 테두리 사각형의 *(R, G, B)*값을 이용하여 보간하게 됩니다. 이때, 각 화소의 색은 화소 중심의 색이고, 각 화소의 중심점 이외의 좌표에서의 색은 주변 중심점 색들을 임의로 혼합한 값이라고 가정합니다. Bicubic 방식을 이용해 확대 영상 화소의 *(R, G, B)*값을 보간하는 방법은 1) 확대 영상 화소의 중심(파란 점)을 기준으로 아래 위 각각 두 줄, 총 네 줄에 대하여 빨간 점으로 표시된 위치에서의 *(R, G, B)*값을 수평 방향 3차식 보간을 통해 얻습니다. 이 수평 방향 3차식 보간은 각 줄에서 빨간 점을 기준으로 좌우 각각 두 개, 총 네 개의 원본 영상 화소값을 이용하여 수행합니다. 2) [그림 3] 좌측의 네 개의 빨간 점에서 *(R, G, B)*값이 구해지면, 이 네 빨간 점을 수직 방향으로 3차식을 보간하여 확대 영상 화소 중심의 *(R, G, B)*값을 구합니다. 그렇다면 이 수평/수직 방향 보간에 사용되는 3차식은 어떻게 알 수 있을까요? 이를 [그림 3] 우측을 통해 설명하겠습니다. [그림 3] 우측은 [그림 3] 좌측의 빨간 점선으로 표시된 각 라인에서 빨간 점 기준으로 좌우 각각 두 개,

조인준
KBS 기술연구소 차장

총 네 개 원본 영상 화소의 R값을 이용해 빨간 점의 R값을 보간하는 경우를 보여주고 있습니다. Bicubic 방식에서 쓰는 3차 보간식에 맞추기 위해 빨간 점 기준으로 좌측 두 화소의 R값 R_{-1}, R_0는 각각 x=-1, 0에, 우측 두 화소의 R값 R_1, R_2는 각각 x=1, 2의 좌표에 대응시키겠습니다. 이렇게 x좌표에 따라 R값을 지정하면, [그림 3] 우측의 파란 점선으로 표시된 완만히 변화하는 함수 R(x)를 가상으로 설정할 수 있고, [그림 2]를 통해 설명된 방식으로 R(x) 함수를 $0 \le x \le 1$ 구간에서 근사화한 [수식 2]의 3차식 함수 f(x)를 구해서 [그림 3] 좌측 빨간 점들에서의 R값을 보간할 수 있습니다. 내용의 올바른 이해를 위해 꼭 기억해야 할 중요 사항은 다음과 같습니다. R(x)는 가상으로 설정된 $0 \le x \le 1$ 구간에서 단순 증가/감소하는 임의의 함수이며 $R(-1)=R_{-1}$, $R(0)=R_0$, $R(1)=R_1$, $R(2)=R_2$를 제외하면 어떤 특정한 식으로 정의되지 않는 대상이라는 것입니다. 이를 꼭 기억하시고, 다음 단계는 R_{-1}, R_0, R_1, R_2값을 이용하여 [그림 3] 좌측 빨간 점에서의 R값을 보간하는 [수식 2]와 같은 3차 보간식 f(x)를 구하는 것입니다. 이를 위해 앞서 설명되었듯 f(x)가 만족해야 하는 조건을 R(x)를 통해 설정합니다.

$$f(0) = R(0),\ f(1) = R(1),\ f'(0) = R'(0),\ f'(1) = R'(1)$$

위 조건에서 f(0)와 f(1)은 $R(0)=R_0$, $R(1)=R_1$를 이용해 쉽게 구할 수 있는데, R(x)가 특정한 식으로 정의되지 않아 R'(x)를 모르는 상황에서 어떻게 R'(0)와 R'(1)을 구할 수 있을까요? 이 문제는 고등학교 시절 수 II 미적분에서 배운 평균값의 정리를 이용합니다. 기억이 나지 않는 독자분들을 위해 평균값의 정리를 다시 한번 상기해보면 [그림 4]와 같이 구간 a, b에서 연속인 임의의 함수 v(x)에 대하여 $\frac{v(b) - v(a)}{b - a} = v'(c)$ 를 만족하는 c(a<c<b)가 적어도 하나 존재한다는 정리입니다.

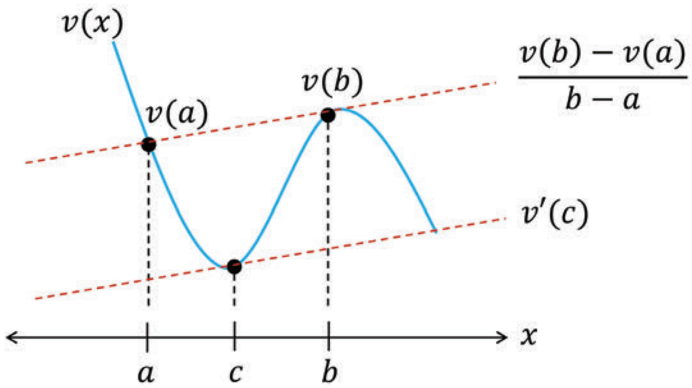

그림 4. 평균값의 정리

이 평균값의 정리를 조금 융통성 있게 적용하여 R'(0)와 R'(1)를 정하면 다음과 같습니다.

$$R'(0) = \frac{R_1 - R_{-1}}{2}$$

$$R'(1) = \frac{R_2 - R_0}{2}$$

이로써 보간을 위한 [수식 2]의 3차식 f(x)가 만족해야 할 조건은 [수식 6]과 같아지고 [수식 4]와 [수식 6]을 이용하여 계산한 f(x)의 계수들은 [수식 7]과 같으며, 이 계수들을 [수식 2]에 대입한 f(x)는 [수식 8]과 같습니다.

[수식 6]

$$f(0) = R_0 \ , \ f(1) = R_1 \ , \ f'(0) = \frac{R_1 - R_{-1}}{2} \ , \ f'(1) = \frac{R_2 - R_0}{2}$$

[수식 7]

$$\begin{aligned} a &= -\frac{1}{2}R_{-1} + \frac{3}{2}R_0 - \frac{3}{2}R_1 + \frac{1}{2}R_2 \\ b &= R_{-1} - \frac{5}{2}R_0 + 2R_1 - \frac{1}{2}R_2 \\ c &= -\frac{1}{2}R_{-1} + \frac{1}{2}R_1 \\ d &= R_0 \end{aligned}$$

[수식 8]

$$\begin{aligned} f(x) = {} & (-\frac{1}{2}R_{-1} + \frac{3}{2}R_0 - \frac{3}{2}R_1 + \frac{1}{2}R_2)x^3 + \\ & (R_{-1} - \frac{5}{2}R_0 + 2R_1 - \frac{1}{2}R_2)x^2 + \\ & (-\frac{1}{2}R_{-1} + \frac{1}{2}R_1)x + R_0 \end{aligned}$$

단, $0 \le x \le 1$

[수식 8]의 f(x)에 [그림 3]에 표시된 빨간 점의 x좌표값 t(t는 모든 확대 영상 화소에 대해 원본 영상과 확대 영상의 가로/세로 화소수의 비율을 통해 계산될 수 있음)를 대입하면 [그림 3] 좌측의 모든 빨간 점에서 R값을 구할 수 있고, 이 네 빨간 점들에서의 R값을 수직 방향으로 같은 방법을 이용하여 보간하면 확대 영상 화소의 R값을 구할 수 있습니다. 이와 같은 방법으로 모든 확대 영상 화소에서 R값, G값, B값을 구하면 Bicubic 방식의 영상 확대가 완료됩니다.

그렇다면 Bicubic 방식이 왜 Bilinear에 비해 좋은 걸까요? 이유는 복잡했던 설명과 달리 다소 허탈할 정도로 단순합니다. Bicubic 방식은 3차식을 이용하여 원본 영상의 4×4 영역을 참조한 보간을 수행하다 보니 1차식을 이용하여 원본 2×2 영상의 영역을 참조하여 보간한 Bilinear 방식에 비해 사람의 눈으로 보기에 확대 영상이 부드러워 보입니다.

지금까지 영상 확대 방법을 다루며 다소 번거로운 과정을 글로 설명하다 보니 읽기에 힘든 부분이 있었을 것 같습니다. 앞으로도 꾸준히 '노력 없이 읽을 수 있는 연재'를 위해 노력하겠다는 약속을 드리며, 다음 연재부터는 영상 확대의 반대 과정인 영상 축소에 관한 설명을 이어가겠습니다.