Week1: Lecture 2: CNN

Vertical edge detection examples

$$\begin{bmatrix} 10 & 10 & 10 & 0 & 0 & 0 \\ 10 & 10 & 10 & 0 & 0 & 0 \\ 10 & 10 & 10 & 0 & 0 & 0 \\ 10 & 10 & 10 & 0 & 0 & 0 \\ 10 & 10 & 10 & 0 & 0 & 0 \\ 10 & 10 & 10 & 0 & 0 & 0 \end{bmatrix} * \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} = \boxed{\begin{bmatrix} 0 & 30 & 30 & 0 \\ 0 & 30 & 30 & 0 \\ 0 & 30 & 30 & 0 \\ 0 & 30 & 30 & 0 \end{bmatrix}}$$

$$\begin{bmatrix} 0 & 0 & 0 & 10 & 10 & 10 \\ 0 & 0 & 0 & 10 & 10 & 10 \\ 0 & 0 & 0 & 10 & 10 & 10 \\ 0 & 0 & 0 & 10 & 10 & 10 \\ 0 & 0 & 0 & 10 & 10 & 10 \\ 0 & 0 & 0 & 10 & 10 & 10 \end{bmatrix} * \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & -30 & -30 & 0 \\ 0 & -30 & -30 & 0 \\ 0 & -30 & -30 & 0 \\ 0 & -30 & -30 & 0 \end{bmatrix}$$

Andrew Ng

Vertical and Horizontal Edge Detection

1	0	-1
1	0	-1
1	0	-1

Vertical

1	1	1
0	0	0
-1	-1	-1

Horizontal

Vertical and Horizontal Edge Detection

1	0	-1
1	0	-1
1	0	-1

Vertical

1	1	1
0	0	0
-1	-1	-1

Horizontal

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
0	0	0	10	10	10
0	0	0	10	10	10
0	0	0	10	10	10

6 x 6

*

1	1	1
0	0	0
-1	-1	-1

=

0	0	0	0
30	10	-10	-30
30	10	-10	-30
0	0	0	0

Andrew Ng

Learning to detect edges

1	0	-1
1	0	-1
1	0	-1

1	0	-1
2	0	-2
1	0	-1

Sobel filter

3	0	-3
10	0	-10
3	0	-3

Scharr filter

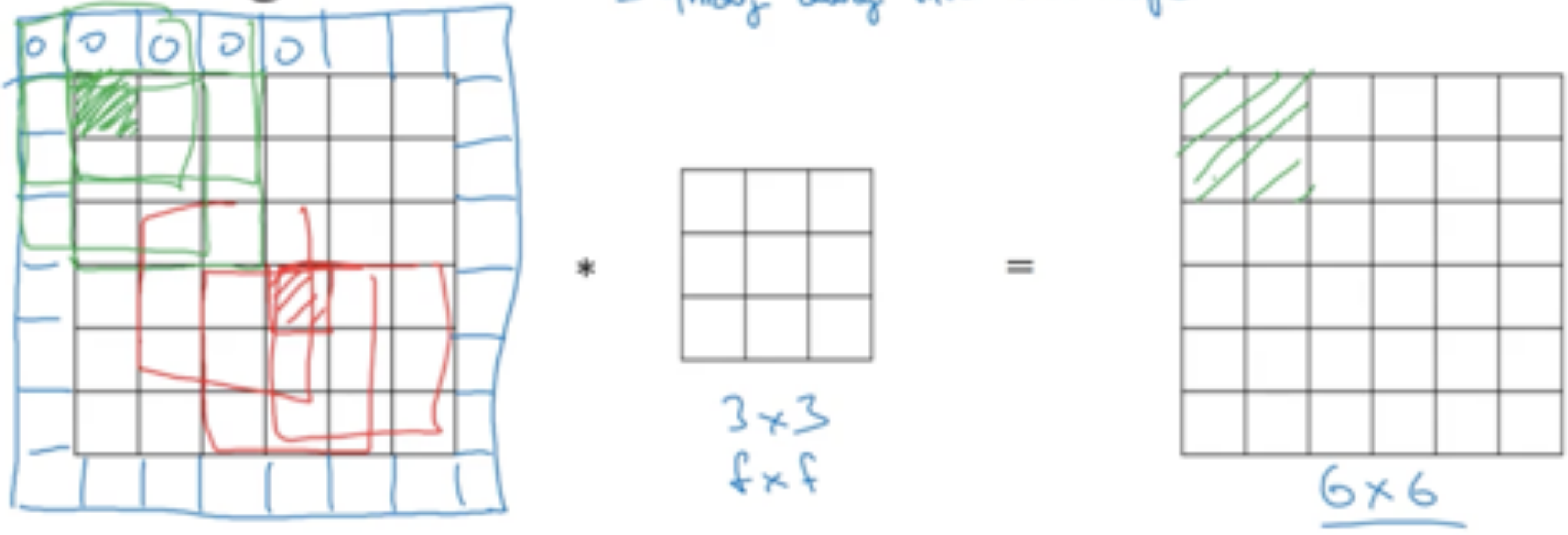
3	0	1	2	7	4
1	5	8	9	3	1
2	7	2	5	1	3
0	1	3	1	7	8
4	2	1	6	2	8
2	4	5	2	3	9

$*$

w_1	w_2	w_3
w_4	w_5	w_6
w_7	w_8	w_9

Padding

- shrinking output
- throw away info from edge



$6 \times 6 \rightarrow 8 \times 8$

$n \times n$

$n-f+1 \times n-f+1$ → ~~4×4~~

$6-3+1=4$

p = padding = 1

$n+2p-f+1 \times n+2p-f+1$

$6+2-3+1 \times \_\_\_ = 6 \times 6$

Andrew Ng

Valid and Same convolutions

no padding

"Valid": $n \times n \quad * \quad f \times f \longrightarrow \underline{n-f+1} \times n-f+1$

$6 \times 6 \quad * \quad 3 \times 3 \longrightarrow 4 \times 4$

"Same": Pad so that output size is the same as the input size.

$n+2p-f+1 \times n+2p-f+1$

f is usually odd

$\cancel{n}+2p-f+1 = \cancel{n} \Rightarrow \underline{p} = \frac{f-1}{2}$

$3 \times 3 \quad p = \frac{3-1}{2} = 1 \quad \Big| \quad \underset{f=5}{5 \times 5} \quad p = 2$

Andrew Ng

7:59 / 9:49

Valid and Same convolutions

no padding

"Valid": $n \times n \quad * \quad f \times f \longrightarrow n-f+1 \times n-f+1$

$6 \times 6 \quad * \quad 3 \times 3 \longrightarrow 4 \times 4$

"Same": Pad so that output size is the same as the input size.

$n+2p-f+1 \times n+2p-f+1$

$\cancel{n}+2p-f+1 = \cancel{n} \Rightarrow \boxed{p = \frac{f-1}{2}}$

$3 \times 3 \quad p = \frac{3-1}{2} = 1 \quad \vert \quad \underset{f=5}{5 \times 5} \quad p=2$

f is usually odd

1×1

3×3

5×5

7×7

Andrew Ng

Strided convolution

2	3	7^{3}	4^{4}	6^{4}	2	9
6	6	9^{1}	8^{0}	7^{2}	4	3
3	4	8^{-1}	3^{0}	8^{3}	9	7
7	8	3	6	6	3	4
4	2	1	8	3	4	6
3	2	4	1	9	8	3
0	1	3	9	2	1	4

7×7

*

3	4	4
1	0	2
-1	0	3

3×3

stride = 2

=

91	100	

Andrew Ng

Strided convolution

2	3	7	4	6	2	9
6	6	9	8	7	4	3
3^{3}	4^{4}	8^{4}	3	8	9	7
7^{1}	8^{0}	3^{2}	6	6	3	4
4^{-1}	2^{0}	1^{3}	8	3	4	6
3	2	4	1	9	8	3
0	1	3	9	2	1	4

7×7

$*$

3	4	4
1	0	2
-1	0	3

3×3

stride = 2

$=$

91	100	83

Strided convolution

2	3	7	4	6	2	9
6	6	9	8	7	4	3
3	4	8	3	8	9	7
7	8	3	6	6	3	4
4	2	1	8	3^{3}	4^{4}	6^{4}
3	2	4	1	9^{1}	8^{0}	3^{2}
0	1	3	9	2^{-1}	1^{0}	4^{3}

7×7

*

3	4	4
1	0	2
-1	0	3

3×3

stride = 2

=

91	100	83
69	91	127
44	72	74

3×3

$n \times n$ * $f \times f$

padding p stride s

$s = 2$

$$\frac{n+2p-f}{s} + 1 \quad \times \quad \frac{n+2p-f}{s} + 1$$

Andrew Ng

Summary of convolutions

$n \times n$ image $\quad f \times f$ filter

padding $p \quad$ stride s

Output size:

$$\left\lfloor \frac{n+2p-f}{s} + 1 \right\rfloor \times \left\lfloor \frac{n+2p-f}{s} + 1 \right\rfloor$$

Technical note on cross-correlation vs. convolution

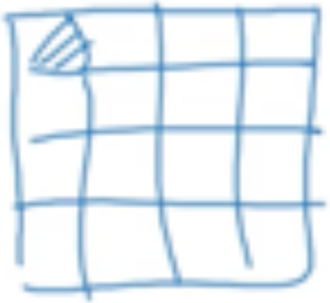
Convolution in math textbook:

2^{7}	3^{2}	7^{5}	4	6	2
6^{9}	6^{0}	9^{4}	8	7	4
3^{-1}	4^{1}	8^{3}	3	8	9
7	8	3	6	6	3
4	2	1	8	3	4
3	2	4	1	9	8

$*$

3	4	5
1	0	2
-1	9	7

=



At 4:44 when Andrew was explaining the technical note on cross-correlation vs convolution, the flipping of the filter was incorrect.

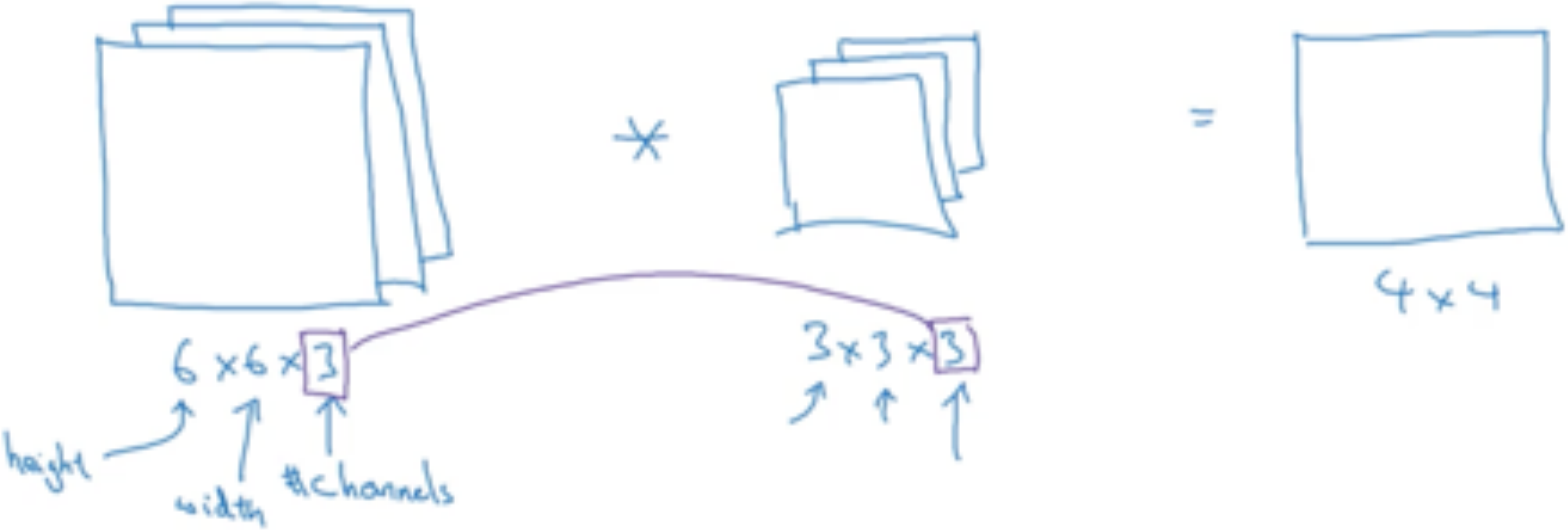
Originally, it was:

3	4	5
1	0	2
-1	9	7

The correct filter after flipping vertically and horizontally would be:

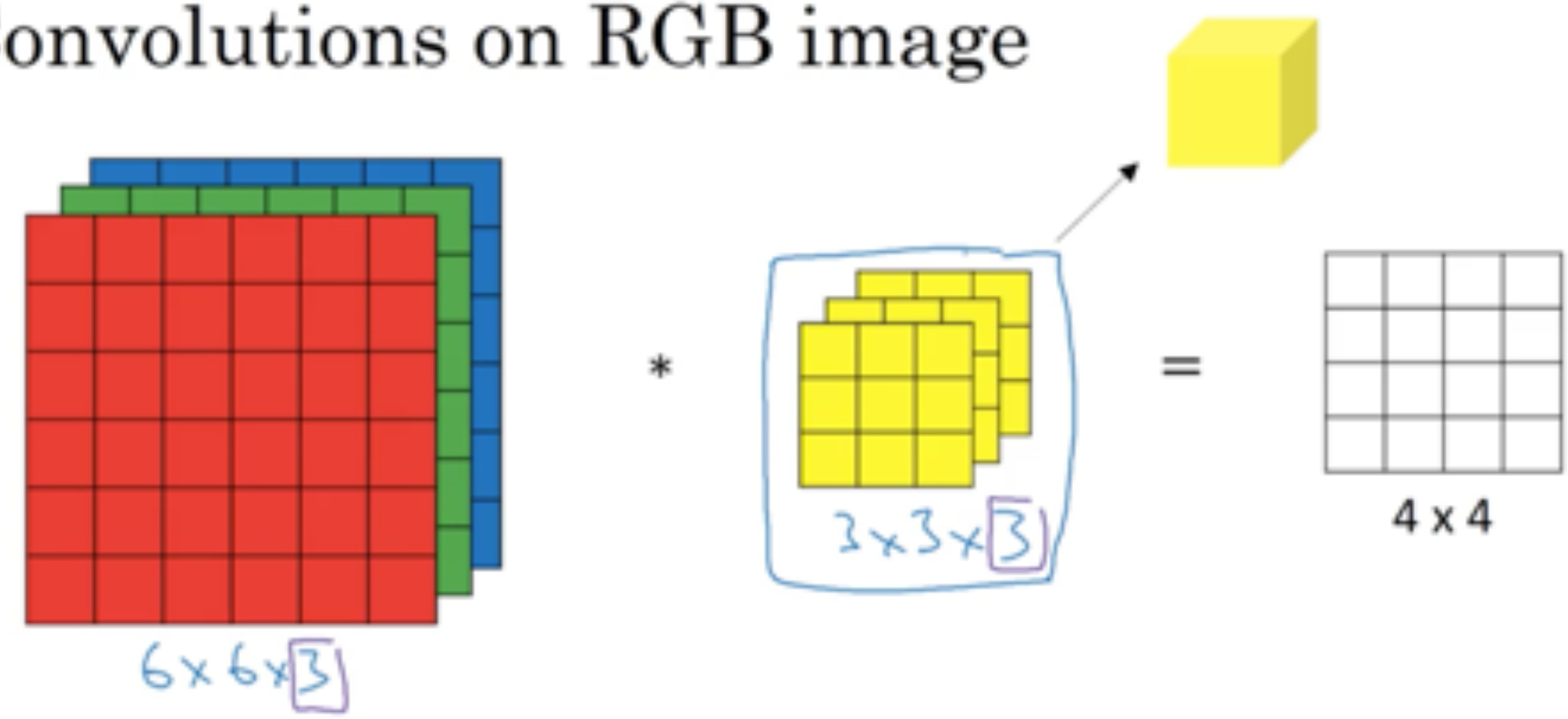
7	9	-1
2	0	1
5	4	3

Convolutions on RGB images



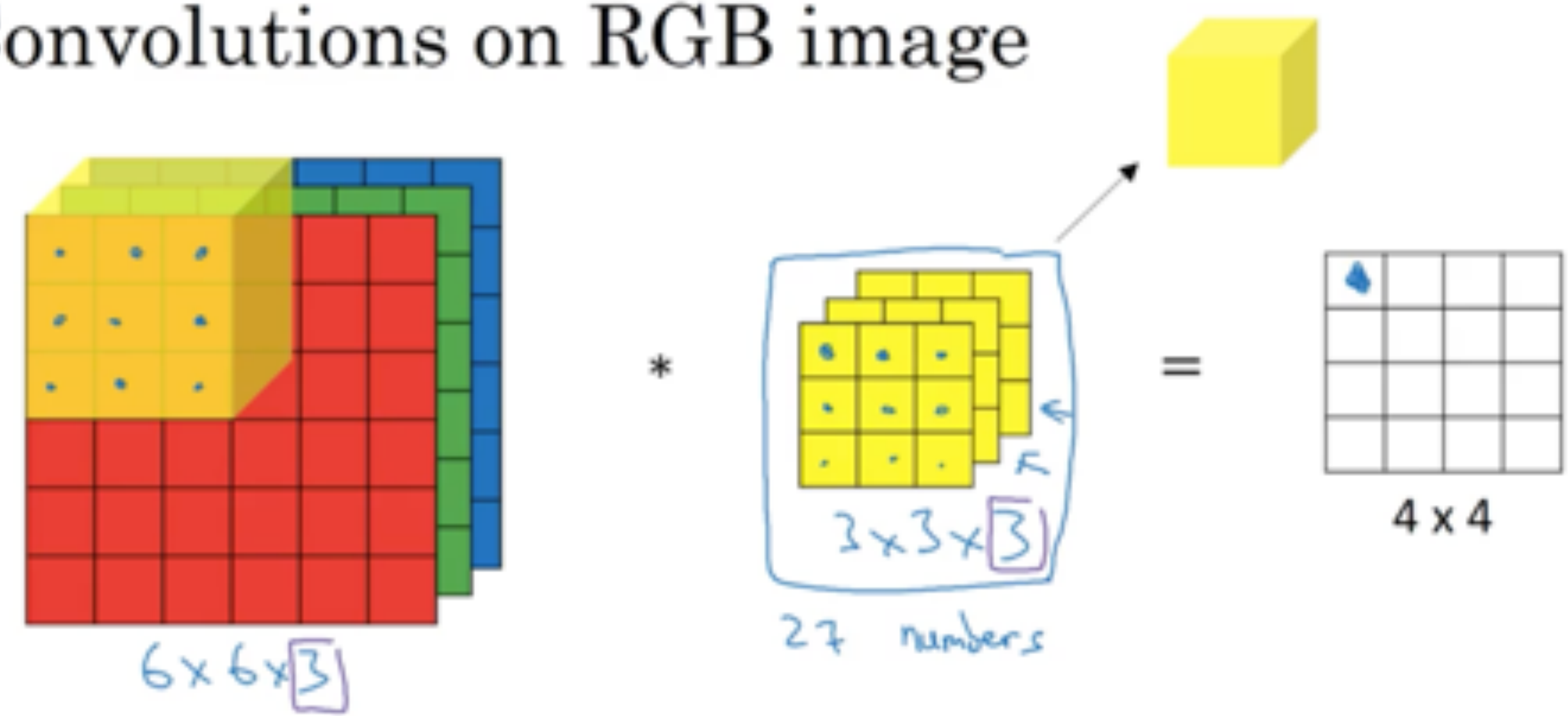
Andrew Ng

Convolutions on RGB image

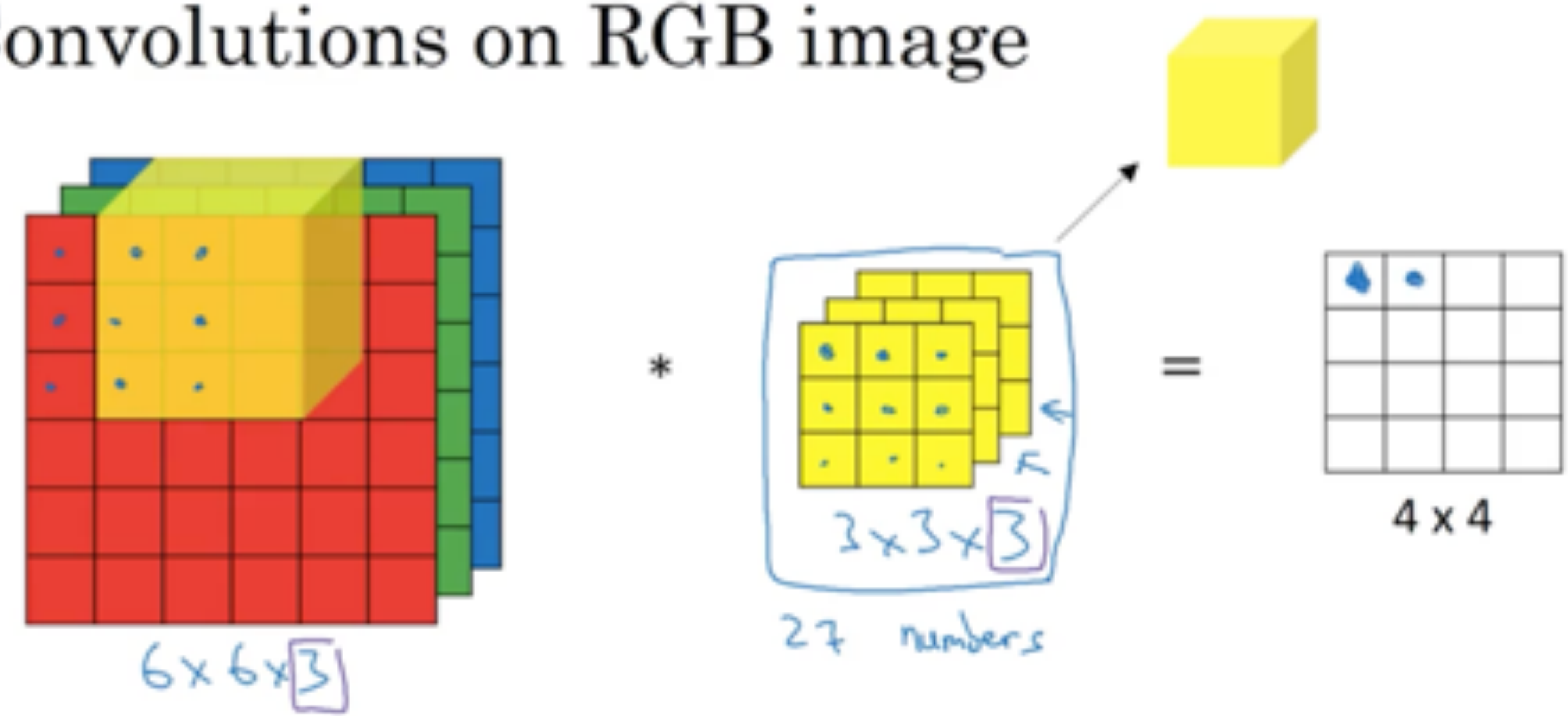


Andrew Ng

Convolutions on RGB image

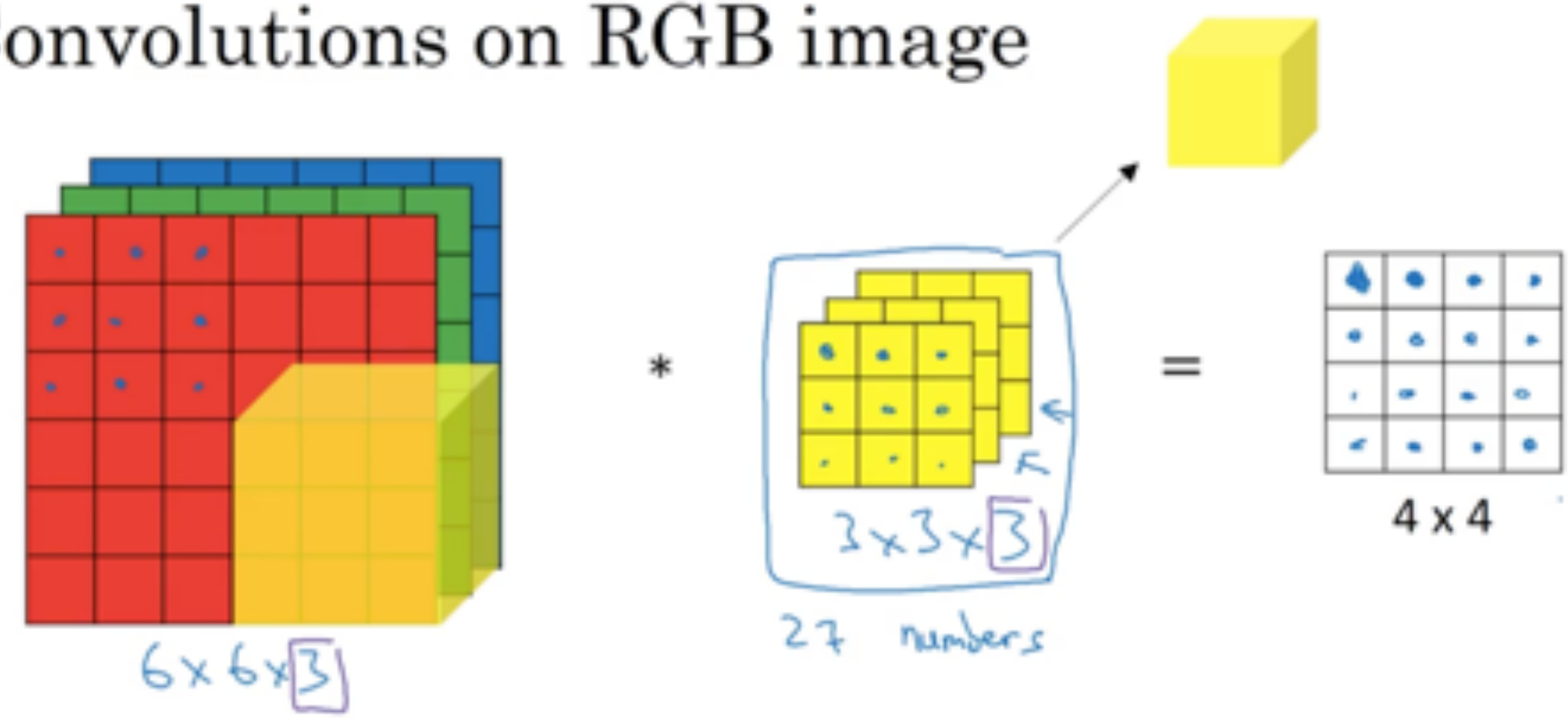


Convolutions on RGB image



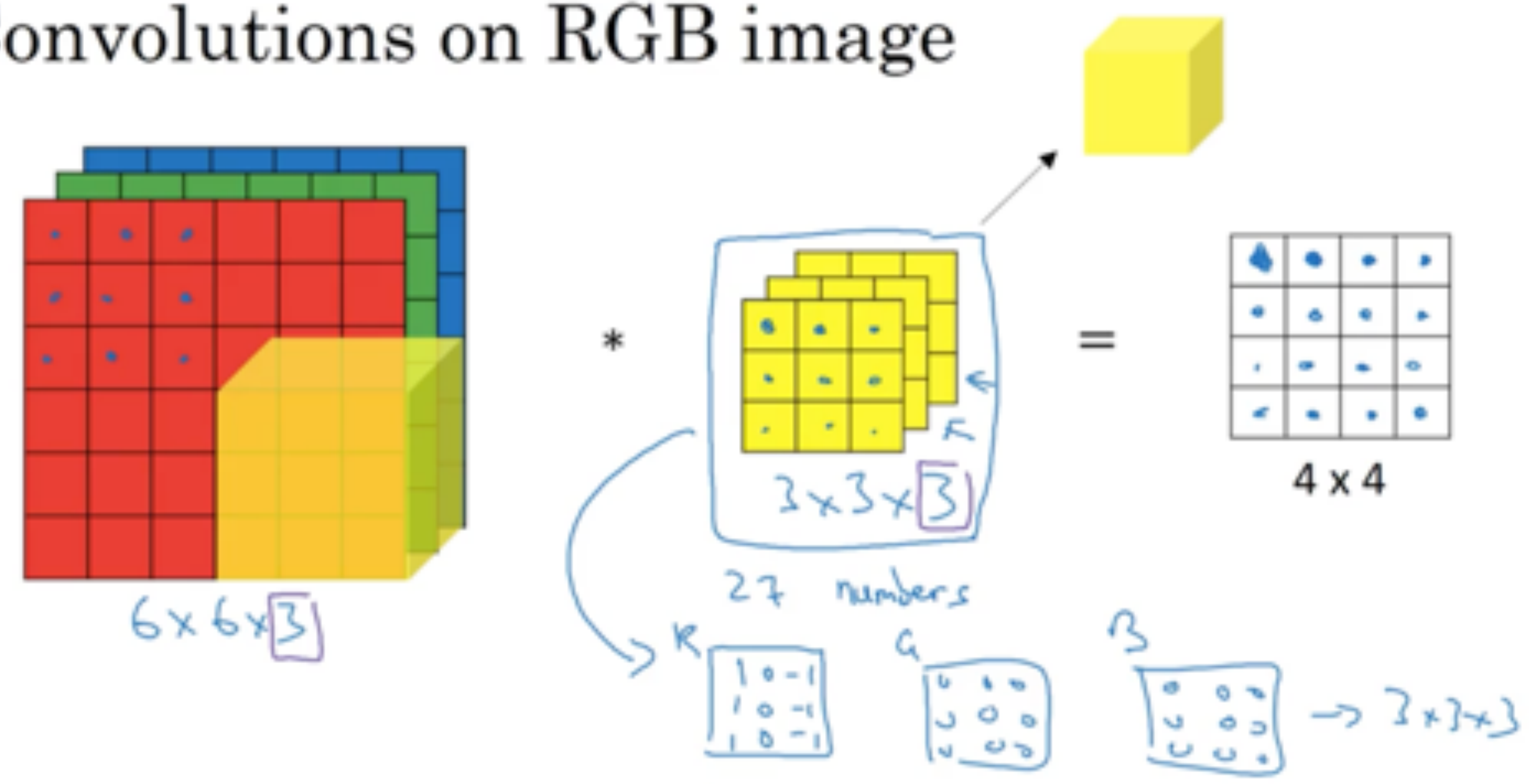
Andrew Ng

Convolutions on RGB image



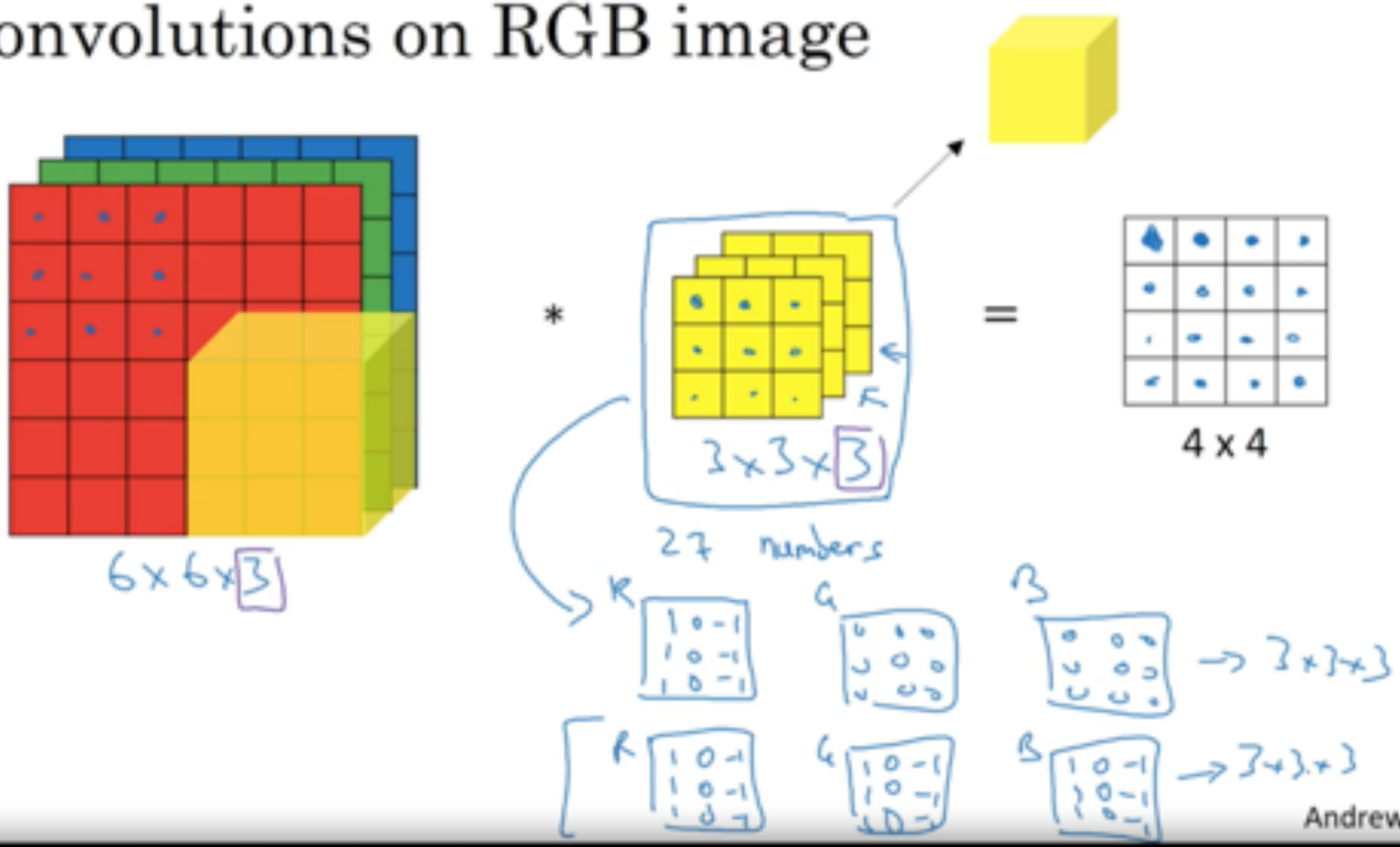
Andrew Ng

Convolutions on RGB image



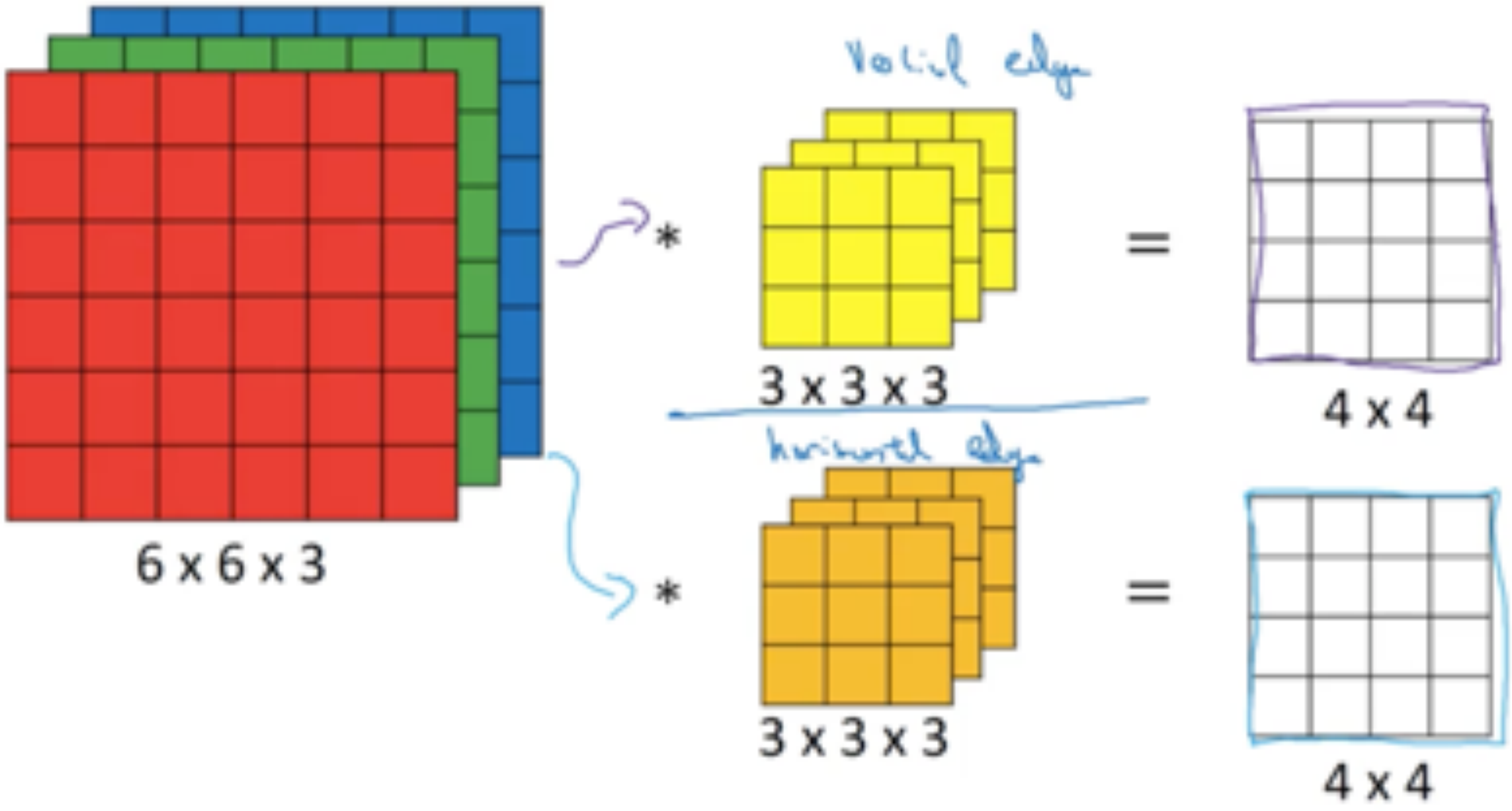
Andrew Ng

Convolutions on RGB image



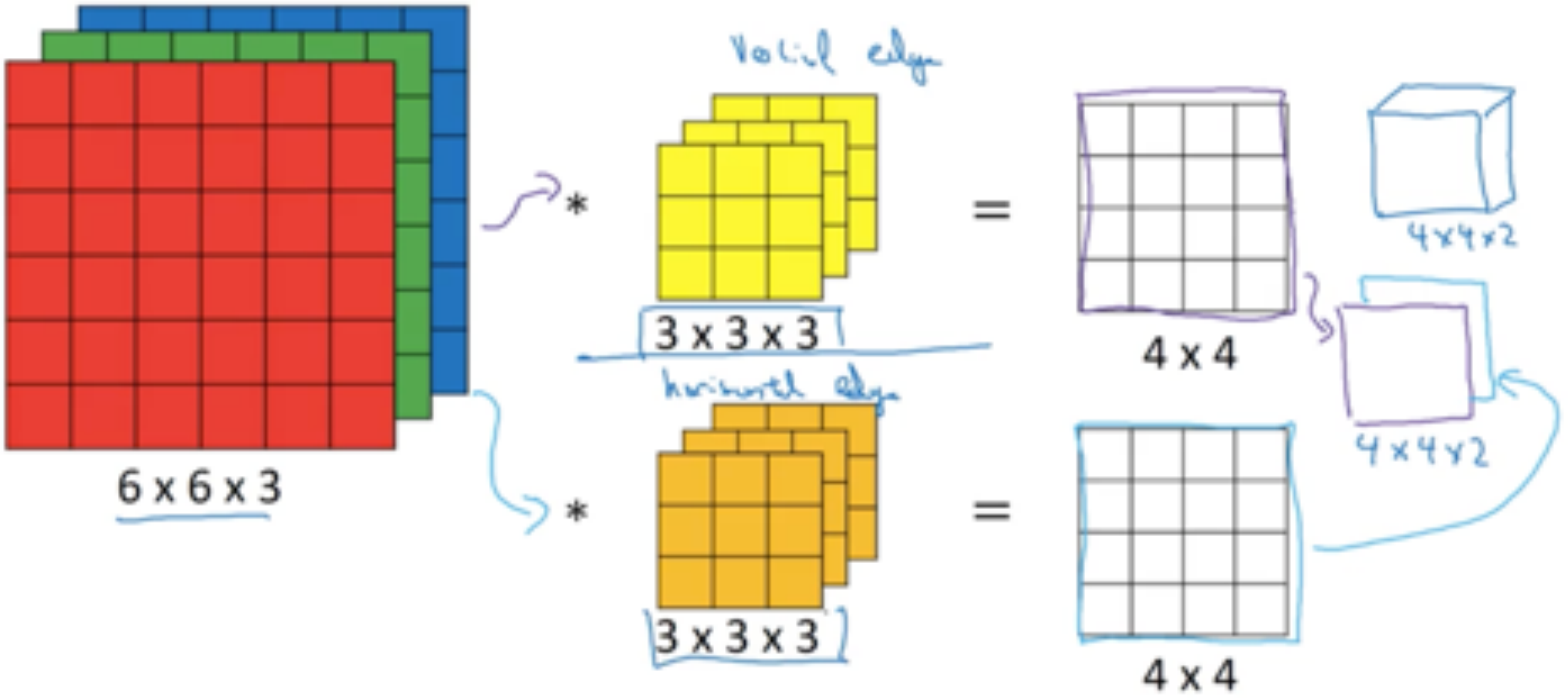
Andrew Ng

Multiple filters



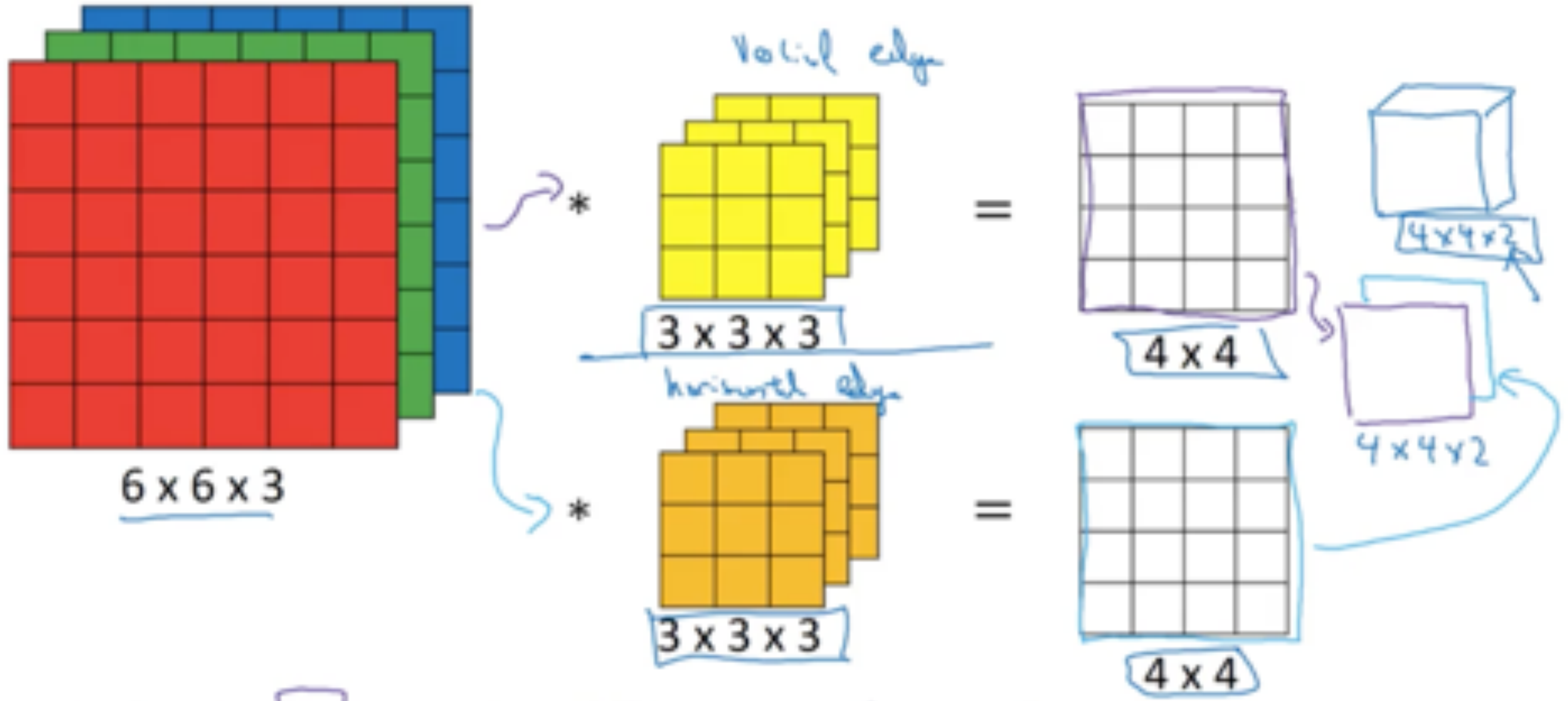
Andrew Ng

Multiple filters



Andrew Ng

Multiple filters



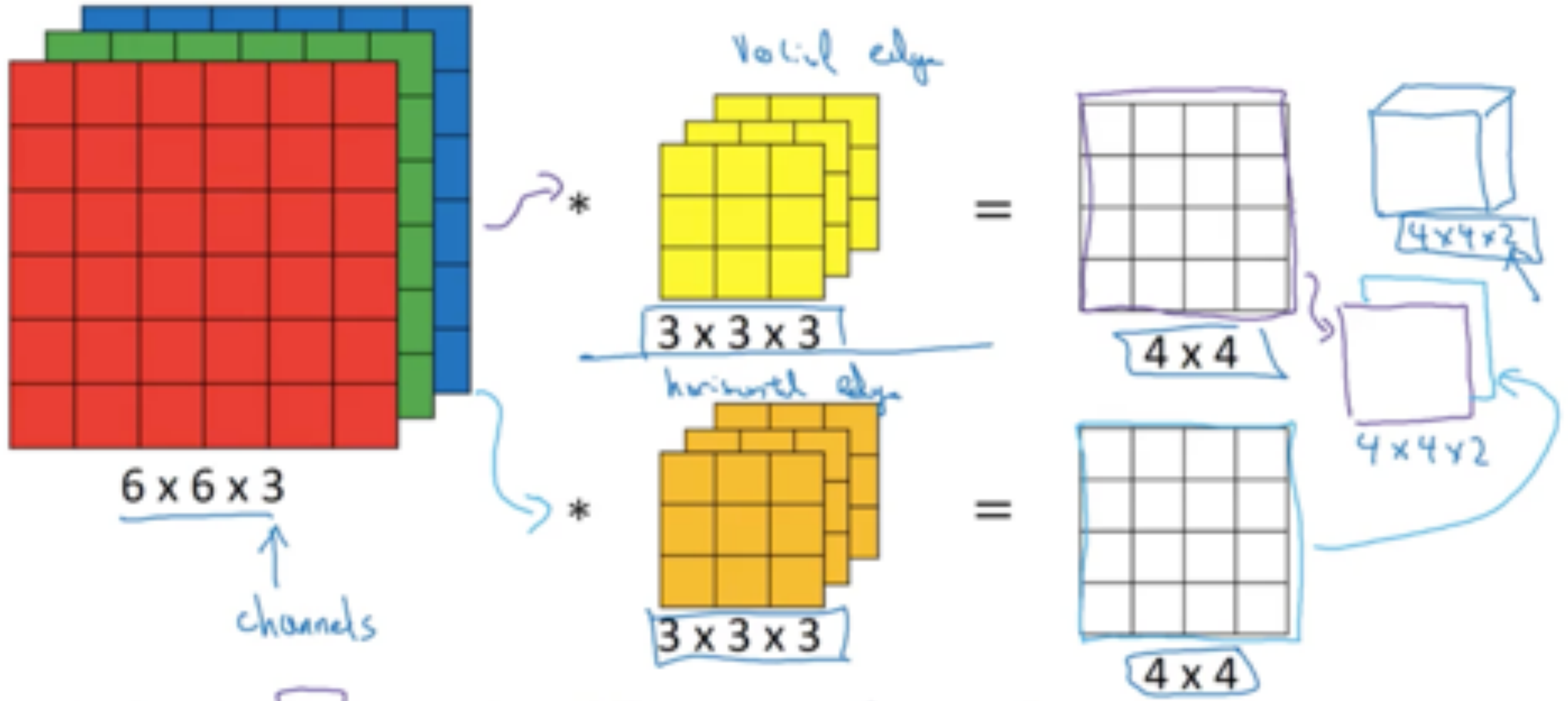
Summary: $n \times n \times \boxed{n_c} \ * \ f \times f \times \boxed{n_c} \ \rightarrow \ \underline{n-f+1} \times \underline{n-f+1} \times n_c'$

$6 \times 6 \times 3 \qquad 3 \times 3 \times 3 \qquad 4 \times 4 \times 2$ ↑ #filters

Andrew Ng

9:17 / 10:44

Multiple filters



Summary: $n \times n \times \boxed{n_c}$ $*$ $f \times f \times \boxed{n_c}$ $\rightarrow$ $\underline{n-f+1} \times \underline{n-f+1} \times n_c'$

$6 \times 6 \times 3$ $\qquad$ $3 \times 3 \times 3$ $\qquad$ $4 \times 4 \times 2$ # filters

Andrew Ng