Motion tracking and cloud height assignment methods for Himawari-8 AMV

Kazuki Shimoji

Japan Meteorological Agency / Meteorological Satellite Center
3-235, Nakakiyoto, Kiyose, Tokyo, Japan

Abstract

Japanese next-generation Himawari-8/9 satellites will carry Advanced Himawari Imager (AHI) units capable of producing full-disk images every 10 minutes with 16 channels. The spatial resolution at the sub satellite point (SSP) is 2 km for IR channels. The use of observation data from these satellites is expected to enable the output of advanced products based on data with high temporal, spatial and spectral resolutions. The Meteorological Satellite Centre of the Japan Meteorological Agency (JMA/MSC) is developing new tracking and height estimation algorithm for Himawari-8 Atmospheric Motion Vector (AMV). The new algorithm is designed for effective utilization of high spatial temporal and spectral resolutions of AHI.

Major changes of the algorithm are applied to cloud feature tracking and cloud height estimation process. In the tracking method, small and large target boxes are prepared respectively for computing two correlation surfaces. Correlation surface from small target box is used as prior information for estimating wind vector, and another correlation surface derived from large target box is used as auxiliary information for determining optimal wind vector which is consistent with both of small and large scale atmospheric motion.

Approach to height estimation method for Himawari-8 AMV is based on optimal estimation to minimize the difference between observed radiance values and the theoretical ones determined from cloud assignment and radiative transfer model parameters using three or more channels. The method will be applied to upper-, medium and low-level clouds for Himawari-8/9 IR/WV wind vectors.

1. INTRODUCTION

In this paper, Basic concepts of Himawari-8 AMV algorithm are provided. Section 2 introduces tracking algorithm based on averaging correlation surfaces, section 3 shows height assignment method based on maximum likelihood estimation. Section 4 provides comparison of statistical characteristics of MTSAT operational IR AMV and MTSAT IR AMV derived by Himawari-8 algorithm. Section 5 introduces comparison of AMV’s heights and CALIPSO. Section 6 describes summary of this paper.
2. TRACKING ALGORITHM

It is considered that using small target box for AMV derivation is good way to retrieve small scale wind. But use of small target box size does not lead to good results necessarily because tracking error is significantly increased. In such case, spurious peak on cross correlation surface can be appeared. This means that information content in very small target box is not enough for pattern matching.

In order to compensate this lack of information, Auxiliary information which can exclude spurious maxima is required. In tracking algorithm of Himawari-8 AMV, averaged surfaces of four correlation surfaces computed in forward and backward matching for small and large target box under assumption that natural atmospheric motion should have temporal and spatial continuity. Way to compute cross correlation itself is the same as MTSAT AMV tracking algorithm. The only difference is that motion vectors for quality control and final output are derived from averaged surfaces.

Cross correlation $C$ for displacement $dx$ and $dy$ computed as follows.

$$
C(dx, dy) \equiv \sum_{x,y} \left( T(x,y) - \frac{1}{NT^2} \sum_{x',y'} T(x',y') \right) \times \left( S(x+dx,y+dy) - \frac{1}{NT^2} \sum_{x',y'} S(x+x'+dx,y+y'+dy) \right)
$$

Where two dimensional arrays $T$ and $S$ corresponds target and search boxes respectively, $NT$ is target box size.

Correlation values on surface is considered as likelihood (or log likelihood) function which represents matching degree between target and search pattern. If correlation surface can be regarded as log likelihood function, the most probable vector temporally and spatially consistent is expected to be derived by finding maximum position of summed likelihood function (averaged correlation surface).

For forward and backward motion vectors for quality control by QI (Holmlund 1998), two averaged correlation surfaces of surfaces by small and large target size ($T_{size1}$ and $T_{size2}$) as follows. Forward motion and backward motion vectors are derived by searching maximum position on these surfaces.

$$
C_{\text{forward}}(dx, dy) \equiv \frac{1}{2} \left( C_{\text{forward}}^{T_{size1}}(dx, dy) + C_{\text{forward}}^{T_{size2}}(dx, dy) \right)
$$

$$
C_{\text{backward}}(dx, dy) \equiv \frac{1}{2} \left( C_{\text{backward}}^{T_{size1}}(dx, dy) + C_{\text{backward}}^{T_{size2}}(dx, dy) \right)
$$

Output motion vector is determined by finding maximum on surface $C_{\text{consistent}}(dx, dy)$, which is described as follows from previously computed $C_{\text{forward}}(dx, dy)$ and $C_{\text{backward}}(dx, dy)$.

$$
C_{\text{consistent}}(dx, dy) \equiv \frac{1}{2} \left( C_{\text{forward}}(dx, dy) + C_{\text{backward}}(-dx, -dy) \right)
$$

Figure 1 shows all of correlation surfaces computed in tracking process. There are too many peaks on not averaged surfaces, but suspicious peaks disappeared after averaging process. It is thought that spurious maxima are mitigated by considering temporal and spatial consistency of natural wind through averaging process regarding correlation as likelihood function. Natural motion vector should satisfy forward and backward matching by small and large target box simultaneously.
In Himawari-8 AMV, three motion vectors are computed from three correlation surfaces. First vector is derived from averaged surfaces by small and large target box in forward matching. Second vector is also same but in backward matching. Those two vectors are used for quality control. Last vector as final output is derived from average of those two averaged surfaces previously computed for first and second vectors. Surface on bottom left is for forward motion, bottom middle is for backward motion and bottom right corresponds to surface for final vector used as output. In reference, averaged surfaces of forward and backward matching to each target box size are shown on right side column. Spurious maxima are mitigated especially in case using small target box.
3. CLOUD HEIGHT ASSIGNMENT METHOD

In IR-WV intercept and histogram analysis methods for JMA heritage AMV algorithm, cloud height is estimated by intensity of radiance and the first guess of humidity and temperature profiles used for atmospheric correction. Namely, heritage algorithm retrieves cloud height by finding optimal cloud height which can simultaneously explain both of observed radiance and NWP first guess without inconsistency. In this scheme, observable information is only radiance.

But by computing motion vector before height assignment process, number of observable could be increased from one to two. Motion vector computed by pattern matching is independent from NWP first guess (if first guess wind is not used for tracking). In Himawari-8 AMV, strategy for cloud height assignment is to search optimal cloud height which can simultaneously explain as many observables (radiance and motion vector) as possible and NWP first guess.

Height assignment method for Himawari-8 AMV is based on maximum likelihood estimation method as same as tracking process. The height assignment consists of five processes.

1. modeling equation for forward model connecting observables and latent variables
2. modeling inequality for constraint for latent variables
3. conversion equation/inequality to likelihood using scaled probability density function and its CDF
4. search optimal latent variables which maximize sum of log likelihood functions
5. select layer corresponding to motion vector

As for first process, radiance forward model which connects latent variables and observables is needed. In Himawari-8 AMV HA algorithm, simple radiance rationing methods using IR bands is used.

Second and third process needs conversion rule from equation and inequality to likelihood function. For example, equation \( f(x) = y \) with error \( \sigma \) is converted to following likelihood function \( L(x, y) \) using Gaussian normal distribution function as follows. (in real application, Cauchy distribution function is thought to be more convenient.)

\[
L_{\text{equation}}(x, y) = \exp\left(-\frac{(f(x) - y)^2}{2\sigma^2}\right)
\]

Inequality \( f(x) > y \) with error \( \sigma \) is converted by using cumulated distribution function \( \text{Erf} \)

\[
L_{\text{constraint}}(x, y) = \frac{1}{2} \left(1 + \text{Erf}\left(\frac{f(x) - y}{\sigma}\right)\right)
\]

For simple example, optimal \( x \) which satisfy both of equation \( x^2 = 4 \) with error \( \sigma_1 \) “AND” constraint \( x > 0 \) with error \( \sigma_2 \), is estimated by maximizing following likelihood function as shown in figure2 (\( \sigma_1 \) and \( \sigma_2 \) are set to 1).

\[
L(x) = L_{\text{equation}}(x) \times L_{\text{constraint}}(x) = \exp\left(-\frac{(x^2 - 4)^2}{2\sigma_1^2}\right) \times \frac{1}{2} \left(1 + \text{Erf}\left(\frac{x - 0}{\sigma_2}\right)\right)
\]

Figure 2: left pannel shows plots for \( L_{\text{equation}}(x) \) (blue) and \( L_{\text{constraint}}(x) \) (yellow). Right panel shows plot for \( L_{\text{equation}}(x) \times L_{\text{constraint}}(x) \) (red). Likelihood for \( x=2 \) (peak around \( x=2 \)) which is the one of two solution for \( x^2=4 \) is mitigated by \( L_{\text{constraint}}(x) \).
For representing complex physical situation, likelihood function values are normalized between 0 and 1 as can be seen as above for fuzzy inference technique which can handle not only "AND" but also "OR" operator.

In fourth process, nonlinear global optimization technique is essential for optimization. Gradient-based optimization technique such as Newtonian method does not work well because there are many local maxima in complex likelihood function and initial point for iteration is not clear. In Himawari-8 AMV, "Differential Evolution method" is used for optimization.

In last process, AMV height is determined by selecting the one of three layers finally retrieved. Basically top layer is assigned as AMV height, but in case that cloud amount on top layer is too small, lower layer is assigned. Threshold to ignore small cloud amount is determined by empirical tuning.

Maximized likelihood function $L_{\text{all}}(\theta_1, p_1, h_1, \theta_2, p_2, h_2, \theta_3, p_3, h_3)$ for retrieving cloud height is represented as follows. The function consists of many parts of likelihood function about radiation, motion consistency between observables and cloud layer structure and constraint to latent variables.

$$L_{\text{all}}(\theta_1, p_1, h_1, \theta_2, p_2, h_2, \theta_3, p_3, h_3) \equiv L_{\text{rad}}(\theta_1, p_1, h_1, \theta_2, p_2, h_2, \theta_3, p_3, h_3) \times L_{\text{wind}}(h_1, h_2, h_3) \times L_{\text{enforcement}}(h, p) \times L_{\text{ordering}}(h_1, h_2) \times L_{\text{ordering}}(h_2, h_3) \times L_{\text{gap}}(h_1, h_2) \times L_{\text{gap}}(h_2, h_3) \times L_{\text{constraint}}(h_1, h_3) \times L_{\text{constraint}}(h_1, h_3) \times L_{\text{constraint}}(h_2, h_3) \times L_{\text{constraint}}(h_2, h_3) \times L_{\text{constraint}}(h_3, h_3)$$

Where $\theta_n, p_n, h_n$ are parameters for cloud amount, phase and height at nth layer.

Likelihood function for radiance consistency $L_{\text{rad}}(\theta_1, p_1, h_1, \theta_2, p_2, h_2, \theta_3, p_3, h_3)$ is defined as follows. The smaller (larger) difference between observed radiance and model radiance, value of this likelihood function draw near to one (zero). The reason that Cauchy distribution is used as a substitute for Gaussian normal distribution is for robustness and preventing underflow in numerical computation.

$$L_{\text{rad}}(\theta_1, p_1, h_1, \theta_2, p_2, h_2, \theta_3, p_3, h_3) \equiv \prod_{ch} \frac{\sigma_{\text{rad}}(ch)^2}{|R_{\text{model}}(\theta_1, p_1, h_1, \theta_2, p_2, h_2, \theta_3, p_3, h_3) - R_{\text{obs}}(ch)|^2 + \sigma_{\text{rad}}(ch)^2}$$

Radiance forward model $R_{\text{model}}(\theta_1, p_1, h_1, \theta_2, p_2, h_2, \theta_3, p_3, h_3)$ is based on radiance rationing method. $\rho_n$ is defined as radiance contribution ratio from nth cloud layer.

$$R_{\text{model}}(\theta_1, p_1, h_1, \theta_2, p_2, h_2, \theta_3, p_3, h_3) \equiv \rho_{g} \epsilon_{g} R_{g} + \rho_{p1} \epsilon(p_1) R(h_1) + \rho_{p2} \epsilon(p_2) R(h_2) + \rho_{p3} \epsilon(p_3) R(h_3)$$

Where $\rho_g$ and $\epsilon_g$ are contribution ratio and emissivity of ground, $\epsilon(p_n)$ is emissivity from nth cloud layer. For simplifying optimization process, $\rho_n$ is described by $\theta$ as follows.

$$\rho_g \equiv (\sin(\theta_1)\sin(\theta_2))^2$$
$$\rho_{p1} \equiv (\sin(\theta_1)\cos(\theta_2))^2$$
$$\rho_{p2} \equiv (\cos(\theta_1)\sin(\theta_3))^2$$
$$\rho_{p3} \equiv (\cos(\theta_1)\cos(\theta_3))^2$$

$\epsilon$ is emissivity of cloud determined by cloud phase parameter $p$. Emissivity values are determined for each bands wavelength.

$$\epsilon(p) = \begin{cases} \text{emissivity of ice cloud}, & 0 \leq p < 1 \\ \text{emissivity of water cloud}, & 1 \leq p < 2 \end{cases}$$

Likelihood function for motion vector is as follows. If only this likelihood function is optimized in single layer case, best fit level is computed. When multiple layers are considered, it is not clear that which layer corresponds to observed motion vector. For solving this problem, fuzzy inference technique is
used for conversion from proposition to likelihood function. Proposition what is needed to convert to likelihood is “observed motion is corresponding to layer #1, #2 OR #3”. This proposition is represented as follow.

\[ P_1 \cup P_2 \cup P_3 \rightarrow P_1 \cap P_2 \cap P_3 \]

Where \( P_1, P_2 \) and \( P_3 \) are

\( P_1 \): observed motion is corresponding to layer #1
\( P_2 \): observed motion is corresponding to layer #2
\( P_3 \): observed motion is corresponding to layer #3

By converting above proposition to membership function (=likelihood function normalized from 0 to 1) using following relationship,

\[ P_1 \cup P_2 \cup P_3 = P_1 \cap P_2 \cap P_3 \rightarrow 1 - (1 - L(P_1))(1 - L(P_2))(1 - L(P_3)) \]

Likelihood for motion vector is represented as

\[ L_{\text{wind}}(h_1, h_2, h_3) \equiv \prod_{ch} \left( 1 - \frac{\sigma_{\text{wind}}(ch)^2}{|u_{FG}(h_n) - u_{obs}(ch)|^2 + |v_{FG}(h_n) - v_{obs}(ch)|^2 + \sigma_{\text{wind}}(ch)^2} \right) \]

Likelihood function to combine motion and radiance information is very important. Role of this function is to prevent over-fitting of motion vector against first guess wind by mandatorily generating cloud as radiation source. “Too convenient” height which has no radiance consistency is excluded by decrease of radiation likelihood function through increase of cloud amount \( \rho \).

The function should means proposition “If difference between motion vector and first guess wind is small at height h, then cloud amount \( \rho \) at h must be greater than \( \rho_{\text{min}} \)” This proposition is represented and converted to likelihood function using following relationship.

\[ P \rightarrow Q = \overline{P} \cap \overline{Q} \rightarrow 1 - L(P)(1 - L(Q)) \]

Where proposition P and Q means

\( P \): difference between motion vector and first guess wind is small at height h
\( Q \): cloud amount \( \rho \) at h must be greater than \( \rho_{\text{min}} \)

Likelihood function for P and Q is as follows.

\[ L(P) : \frac{\sigma_{\text{wind}}^2}{|u_{FG}(h) - u_{obs}|^2 + |v_{FG}(h) - v_{obs}|^2 + \sigma_{\text{wind}}^2} \]

\[ L(Q) : \frac{1}{\pi} \arctan \left( \frac{\rho - \rho_{\text{min}}(ch)}{\sigma_q} \right) + \frac{1}{2} \]

As a result, likelihood function for proposition “If difference between motion vector and first guess wind is small at height h, then cloud amount \( \rho \) at h must be greater than \( \rho_{\text{min}} \)” is represented as

\[ L_{\text{cloud}}(h, \rho) \equiv \prod_{ch} \left( 1 - \frac{\sigma_{\text{wind}}(ch)^2}{|u_{FG}(h) - u_{obs}(ch)|^2 + |v_{FG}(h) - v_{obs}(ch)|^2 + \sigma_{\text{wind}}(ch)^2} \right) \times \left( 1 - \frac{1}{\pi} \arctan \left( \frac{\rho - \rho_{\text{min}}(ch)}{\sigma_q} \right) + \frac{1}{2} \right) \]
In addition, for introducing layer ordering and setting gap between layers, following functions are used. Setting gap between layers is introduced for preventing radiance over-fitting.

\[ L_{\text{constraint}}^{\text{ordering}}(h_i, h_j) \equiv \frac{1}{\pi} \arctan\left(\frac{h_i - h_j}{\sigma_{\text{height}}}\right) + \frac{1}{2} \]

\[ L_{\text{constraint}}^{\text{gap}}(h_i, h_j) \equiv \frac{1}{\pi} \arctan\left(\frac{h_i - h_j - \Delta h}{\sigma_{\text{height}}}\right) + \frac{1}{2} \]

4. STATISTICS AGAINST SONDE AND FG WIND

JMA/MSC compared operational MTSAT AMVs (RTN) and new AMVs (TEST) retrieved by new tracking and height assignment methods. 5x5 and 15x15 pixels target box is utilized in new tracking method. Observed brightness temperature of IR1 (10.8 micro meter), IR2 (12.0 micro meter) and IR3 (6.7 micro meter) bands and motion vector computed from IR1 and IR3 imagery are used for HA. 15 minutes separated imagery is used as same as operational 6-hourly MTSAT AMV. Periods of this experiment is for January 2013. In winter season of East Asia region, strong negative wind speed BIAS is observed in operational AMV due to Jet stream as shown in figure 4.

Table 1 shows comparison of IR upper level AMVs derived by operational AMV (left) and by new tracking and height assignment methods (right). The most significant difference is the average speed. Totally new tracking algorithm tends to compute higher wind speed than operational wind. As for quality, RMSVD are decreased over northern hemisphere, neutral over tropical region and slightly debased at medium level over southern hemisphere.

Figure 2 shows verification of RMSVD (O-B) map. Large RMSVD around jet stream region over northern hemisphere are mitigated significantly. This mitigation can be confirmed around upper level (100-400hpa) as seen in Figure 3. Characteristic of wind speed BIAS is shown in Figure 4 and 5. Negative speed bias decreased over northern hemisphere, and two-dimensional and vertical coherence of wind speed BIAS is improved. But totally positive wind speed BIAS can be seen as seen in Figure 5.

Table 1: comparison of AMV vs sonde statistics with current operational AMV(left) and by Himawari-8 AMV algorithm (right).
Figure 3: Verification of RMSVD (O-B) with the current operational MTSAT-AMV algorithm (left) and the new MTSAT AMV by Himawari-8 AMV algorithm using only radiance information (middle) and using both of radiance and motion information (right) for January 2013.

Figure 4: As per Figure 3, but in vertical statistic. Red, orange, green and blue corresponds to fulldisk, northern hemisphere (N20-N60), tropical region (N20-S20) and southern hemisphere (S20-S60) respectively.

Figure 5: Verification of RMSVD (O-B) with the current operational MTSAT-AMV algorithm (left) and the new MTSAT AMV designed for Himawari-8 using only radiance information (middle) and using both of radiance and motion information (right) for January 2013.
5. COMPARISON BETWEEN CALIPSO AND COLLOCATED AMV HEIGHT

Figure 6 and 7 is comparison of AMV height retrieved by Himawari-8 algorithm and CALIPSO back scatter plot for winter and summer season. Generally cloud layer structure in winter season over Asia region is not so complicated as seen in Figure 6. As a result, layer selection for AMV height is not difficult. This simplicity may be one of the reason that RMSVD and wind speed BIAS is improved in winter hemisphere.

But seeing Figure 7 which is for summer season, multiple layer structure can be found. The multiple layer structure makes layer selection process difficult. Parameter tuning for ignoring very thin cirrus is thought to be essential for improvement to height assignment accuracy.

**Figure 6**: As per Figure 4, but in vertical statistic. Red, orange, green and blue corresponds to full disk, northern hemisphere (N20-N60), tropical region (N20-S20) and southern hemisphere (S20-S60) respectively.

**Figure 7**: upper panel is collocation plot of IR AMV height (red dots) computed by new HA method and calipso backscatter coefficient plot for 18UTC 18th February 2014. Bottom panels show CALIPSO path and map of the collocated AMVs and MTSAT imagery.
6. SUMMARY

Japanese next-generation Himawari-8/9 satellites has a capability to observe full-disk images every 10 minutes with 16 channels. The spatial resolution at SSP will be upgraded 2 km for IR channels. For effective use of temporal and spatial resolution of Himawari-8, JMA/MSC has developed tracking and height assignment algorithm. Both of tracking and height assignment methods are based on maximum likelihood estimation (MLE) method. As for tracking, MLE is introduced by regarding correlation surface as log likelihood function and motion vector is derived from average of correlation surfaces computed from small and large target box with forward and backward matching. Cloud height are estimated by finding optimal cloud structure which can explain both of observed radiances and motion vectors by multiple bands. Statistical result for MTSAT AMV derived from new tracking and height assignment shows improvement to negative wind speed BIAS in winter season over northern hemisphere. But positive wind speed BIAS can be found in summer hemisphere. By comparing AMV height with CALIPSO product, reason of the positive BIAS in summer season could be height selection error caused from multi layer situation. It is thought that studying how to determine optimal threshold to ignore thin cirrus cloud is important.
7. REFERENCES

Christopher M. Bishop, 2006: Pattern Recognition and Machine Learning, Springer


Cotton, J., Forsythe, M., 2010: Fourth analysis of the data displayed on the NWP SAF AMV monitoring

Shimoji, K., 2010: The Development for MTSAT Rapid Scan High Resolution AMVs at JMA/MSC, Proceedings of the Tenth International Winds
