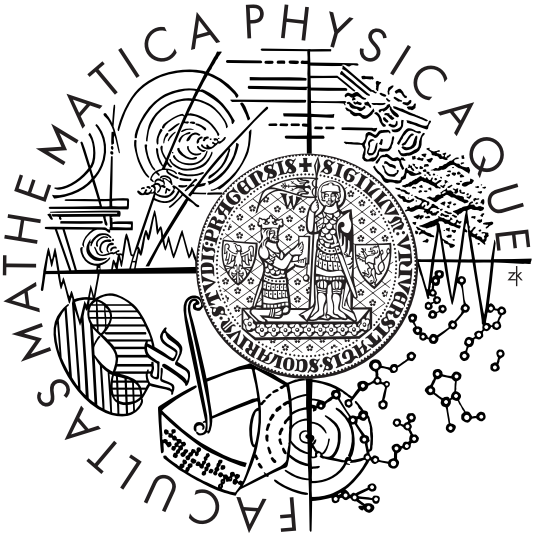
Charles University in Prague

Faculty of Mathematics and Physics

MASTER THESIS



Petra Vaníčková

Analysis of meteorological images for storm detection in short-term weather forecast

Institution: Department of software engineering

Supervisor: Ing. Filip Šroubek, Ph.D.

Study branch: Computer graphics

I would like to express my sincere gratitude to my supervisor Ing. Filip Šroubek, Ph.D., who patiently guided this study and his advice significantly improved quality of this thesis.

I gratefully acknowledge RNDr. Martin Setvák, CSc for providing me with both the thesis subject and all necessary data. His help during the testing phase of the study was invaluable.

Prohlašuji, že jsem svou diplomovou práci napsala samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

V Praze dne 18.4.2007 Petra Vaníčková

Contents

Problem definition

V obrazových datech z přístroje SEVIRI (Spinning Enhanced Visible and Infrared Imagery) družic MSG (Meteosat Second Generation) lze vyhledávat struktury označované jako "studené-U" (resp. "studené-V"), které se občas vyskytují v rámci horní hranice oblačnosti konvektivních bouří. Tyto struktury, detekovatelné v tepelných kanálech meteorologických družic, se většinou vyskytují v souvislosti s extrémními projevy daných bouří, a proto se považují za jeden z možných příznaků potenciální nebezpečnosti bouří, využitelný pro velmi krátkodobou předpověď počasí. Úkolem řešitele je navrhnout automatickou (poloautomatickou) metodu, která by nalezla struktury typu „U“, dle následujícího postupu: - stručný popis jevu (na základě vybraných článku v odborné literatuře) - teoretický návrh řešení - zhodnocení dosavadních výzkumů a aplikací na podobných úlohách - navržení vlastní metody pro klasifikace bouří s využitím trénovací množiny - implementace návrhu - ověření úspěšnosti klasifikace na reálných datech

The imagery obtained from the SEVIRI (Spinning Enhanced Visible and Infrared Imagery) instrument of the satellites MSG (Meteosat Second Generation) can be used for searching for structures denoted "cold-U" ("cold-V" respectively). These structures are occasionally present in upper parts of the convective storm clouds. These structures, which can be detected in radiosity chanells of the meteorological satellites, are mostly connected to the severe storms, and thus they are considered to be a possible sign of a storm severity usable for nowcasting. The aim of the study is to propose an automatic (semiautomatic) method which would be able to detect the "cold-U" occurrences.
The study should contain the following steps:
- brief description of the feature (based on the chosen technical papers) - theoretical proposal of a solution - comparison with existing approaches and applications - proposal of a concrete classification algorithm with using a training data set - implementation of the algorithm - verification of the algorithm accuracy on a testing data set

Abstract

Název práce: Analýza meteorologických snímků za účelem detekce bouří pro velmi krátkodobou předpověď počasí
Autor: Petra Vaníčková
Obor: Softwarové systémy
Vedoucí diplomové práce: Ing. Filip Šroubek, Ph.D.
E-mail vedoucího: sroubekf@utia.cas.cz
Abstrakt: Meteorologické satelitní systémy prodělaly v průběhu předchozích patnácti let rychlý vývoj. Došlo nejen ke zvýšení frekvence pořizovaných snímků, ale i ke zlepšení jejich rozlišení. Tento technologický posun otevírá nové možnosti k rozsáhlejšímu využití družicových snímků pro předpověď počasí. Cílem této studie je navržení algoritmu, který by automaticky nacházel jev nazývaný "enhanced-U/V". Významnost tohoto jevu spočívá ve faktu, že jeho výskyt je vysoce korelovaný s výskytem silných bouří. Algoritmus používá obrazový materiál získaný ze specifického typu infračerveného kanálu geostacionární družice. Navrhované řešení zahrnuje stručný popis meteorologické motivace, převedení zadání na informatický problém, návrhy možných přístupů k řešení, diskuzi nad výsledky získanými v průběhu vývoje a testování prototypového řešení a shrnutí dalších možných vylepšení zvoleného postupu.Výsledky testů ukazují, že algoritmus je poměrně úspěšný ve vyhledávání "enhanced-U/V" označených expertem, na druhou stranu je stále příliš vysoký počet oblastí nesprávně označených za "enhanced-U/V". Toto chování je možné vylepšit několika postupy, např. sledováním historického vývoje podezřelé lokality v časové posloupnosti snímků.
Klíčová slova: Zpracování obrazu, Předpověď počasí, Detekce bouří

Title: Analysis of meteorological images for storm detection in short-term weather forecast
Author: Petra Vaníčková
Branch of study: Software Systems
Supervisor: Ing. Filip Šroubek, Ph.D.
Supervisor's email address: sroubekf@utia.cas.cz
Abstract: Meteorological satellite systems have been developing rapidly during the last fifteen years and both imagery resolution and shot frequency have been improved. This evolution opens possibilities for more extensive usage of the satellite imagery in weather forecasting. The purpose of this thesis is to design an algorithm which would automatically detect the enhanced-U/V features. The importance of the enhanced-U/V feature lies in the fact that it is highly correlated with occurrences of severe storms. The algorithm uses images acquired from a specific IR channel of a geostationary satellite. The proposed solution includes a brief introduction to the problem from the meteorological point of view, an analysis of the task as a problem of the computer science, a proposal of possible approaches to the problem, a discussion of the results obtained during the development and testing of the prototype software, and a summary of suggested algorithm improvements.The results of the tests show that the algorithm is rather successful in seeking the enhanced-U/V (marked by an expert) but on the other hand the number of false alarm cases is still too high. This behaviour can be improved by several techniques e.g. considering the case's history from series of images.
Keywords: Image processing, Weather forecasting, Storm detection

Chapter 1

Introduction

A significant progress in quality of the sattelite imagery has been made during the last fifteen years. The improvement of the image resolution and the higher acquisition frequency have also increased the amount and quality of the data extracted from various meteorological sattelite systems. Therefore, the sattelite imagery can be more extensively used as a forecasting tool.

One of the features which can be recognized in the IR sattelite images is the enhanced-U/V (also called cold-U or cold-V). Each occurrence of the feature implies a high probability of a severe storm (e.g. hails, strong wind or tornado). This correlation makes the enhanced-U/V an object of interest. Although this structure has been described in many meteorological works, its detection still depends on the examination by an expert and no attempts at the automatic detection have been implemented before this study.

This study proposes a novel approach to the automatic enhanced-U/V detection based on standard techniques of the image processing science. A pilot application, which is designed according to the proposal, is supposed to verify the suitability of used methods and to disclose eventual weaknesses of the algorithm. The ability of the enhanced-U/V detection is consequently evaluated by a test battery. The algorithm proposal is restricted to a single image analysis, however, a part of the testing phase investigates a possible adaptation of the algorithm so it would obtain additional information from a series of images.

The first part of this thesis introduces the enhanced-U/V feature and its meteorological background. Physical properties of the structure are described and several possible complications caused either by the scanning technique or by the uneven atmospherical structure are mentioned. The second chapter presents the algorithm proposal and transforms the meteorological motivation to a problem of the image processing. Three basic components of the algorithm are described - segmentation, feature extraction and classification. The third chapter is dedicated to the tests and discussion of the results. The fourth chapter summarizes future works and possible algorithm improvements which ensue both from the tests and from further observed properties of the enhanced-U/V.

Chapter 2

Meteorological background

This chapter focuses on the physical motivation of the study. The following text describes basic facts about the sought feature in the context of meteorology and specifies character of the satellite imagary, which the application uses to detect enhanced U/V shapes in.
For better understanding of the following sections some terms need to be clarified, the following definitions have been taken from .[4].

An expression **convective-clouds** describes the vertically developed family of clouds, most often initiated or forced by atmosperic instability, convection. The height of their bases ranges from as low as 300 m up to about 3000m. Their height ranges from tens of meters (shallow convective clouds) up to 10 to 20 km, (deep convective clouds), depending on height of the tropopause and updraft strength. Clouds with extensive vertical development are indication of unstable air. Strong upward currents (updrafts) in vertically developed clouds can carry high concentrations of supercooled water to high levels where temperatures are quite cold. Upper portions of these clouds may be composed of mixture of supercooled water and ice, while their cloud tops are typically composed of ice.
Typical representative of this group is **Cumulonimbus cloud** which is characterized by strong vertical development in the form of mountains or huge towers topped at least partially by a smooth, flat, often fibrous anvil. Also known colloquially as a "thunderhead."

Equivalent black body temperature (BT) is the surface (skin) temperature which the body would have if it were in radiative equilibrium and would emmit according to Planck equation, with emissivity equal to one and with no atmospheric absorption/emission above the emitting surface, typically expressed , in Kelvin [K].

By the term **storm** is meant any disturbed state of the atmosphere, especially affecting the Earth's surface, and strongly implying destructive and otherwise unpleasant weather. Storms range in scale from tornadoes and thunderstorms to tropical cyclones to synoptic-scale extratropical cyclones. This work mainly focuses on the severe type of convective storms. **Severe convective storms** threaten life and/or property. Examples include severe thunderstorms with large hail, damaging wind, or tornadoes. Although cloud-to-ground lightning is not among criteria for classifying a storm as

severe, it is perceived as a highly dangerous phenomenon, a leading cause of deaths, injuries, and damage from thunderstorms. A thunderstorm needs not be severe to generate frequent cloud-to-ground lightning. Additionally, excessive localized convective rain is not classified as severe storm but often accompanies severe storms. Such rainfall may result in related phenomena (flash floods) that threaten life and property.

A supercell is relatively long-lived thunderstorm with a persistent rotating updraft. Supercells are rare, but are responsible for a remarkably high percentage of severe weather events - especially tornadoes, extremely large hail and damaging straight-line winds. They frequently travel to the right of the main environmental winds (i.e. they are right movers). Radar characteristics often (but not always) include mesocyclone. **Mesocyclone** is a storm-scale region of rotation, typically around 3-10 km in diameter and often found on the right rear flank of a supercell. The circulation of a mesocyclone covers an area much larger than the tornado that may develop within it. Properly used, mesocyclone is a radar term; it is defined as a rotation signature appearing on a Doppler radar that meets specific criteria for magnitude, vertical depth, and duration. Therefore, a mesocyclone should not be considered a visually-observable phenomenon (although visual evidence of rotation, such as curved inflow bands, may imply the presence of a mesocyclone).

Another part of a storm cloud is **the anvil** which is the flat, spreading top of a cumulonimbus cloud, often shaped like an anvil. Thunderstorm anvils may spread hundreds of miles downwind from the thunderstorm itself, and sometimes may spread upwind.

The parts of storm clouds which can reach the highest altitudes are **overshooting tops** (or penetrating towers) - dome-like protrusions above a thunderstorm anvil, representing a very strong updraft and hence a higher potential for severe weather with that storm. A persistent and/or large overshooting top (anvil dome) often is present on a supercell.

Figure 2.1displays a typical shape of a storm cloud, more specifically of a supercell.

Next few paragraphs briefly describe basic layers of the atmosphere.
Troposphere is the layer of the atmosphere from the Earth's surface up to the tropopause, characterized by decreasing temperature with height, vertical wind motion, appreciable water vapor content, and sensible weather (clouds, rain, etc.).
Tropopause is the upper boundary of the troposphere, usually characterized by an abrupt change in lapse rate from positive (decreasing temperature with height) to neutral or negative (temperature constant or increasing with height).
Stratosphere is the region of the atmosphere extending from the top of the troposphere to the base of the mesosphere (up to about 50 km), an important layer containing stratospheric ozone.
Mesosphere is the atmospheric shell between about 50 km and about 70 to 80 km, extending from the top of the stratosphere (the stratopause) to the upper temperature

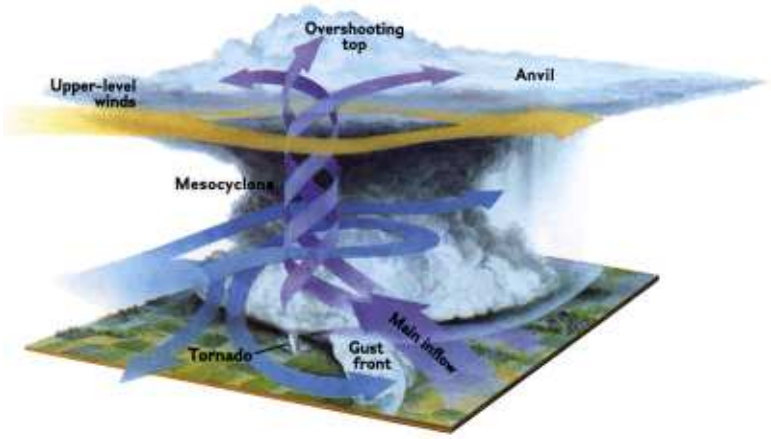


Figure 2.1: Scheme of (tornadic) supercell

minimum that defines the mesopause (the base of the thermosphere).
Mesopause is the top of the mesosphere, corresponding to the level of minimum temperature in the atmosphere found at 70 to 80 km.
Atmospheric layers are shown in Figure 2.2.

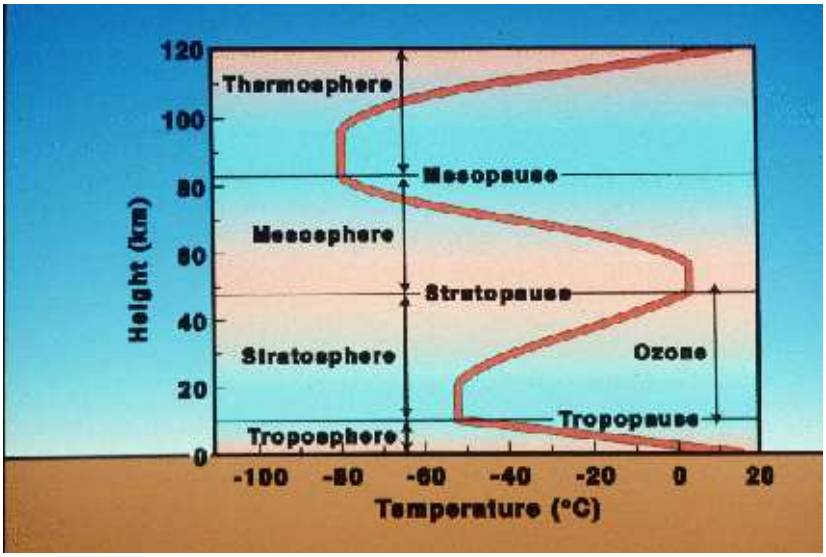


Figure 2.2: Atmospheric structure [19]

2.1 Definition of the enhanced U/V feature

Since the 1980's many scientific works have analyzed the possibility of the severe storm detection from the data obtained from weather satellites. One of the recognized features, which can be observed at the satellite images in emissive bands, is called enhanced-U/V feature. Presence of the enhanced-U/V feature is closely connected with occurrences of severe storms. As has been documented, about 70% of storms containing enhanced-U/V are severe.[15] On the other hand approximately 45% of severe storms do not contain enhanced-U/V and that's why this method cannot be the only tool for the severe storms nowcasting (a very short term forecasting). Median lead time to the first onset of the enhanced-U/V to the first report of severe weather is approximately 30 minutes.[15]
The goal of this study is to develop an algorithm which would automatically detect parts of clouds which are suspicious of being enhanced-U/V.

The enhanced-V (or more precisely „cold-U/V“) term referes to a feature observed occasionally in thermal IR wavelengths (at any of the emissive satellite bands, typically but not exclusivelly at 10 – 12.5 μm) at tops of deep convective storms, resembling a cold U or V shape, surrounding an enclosed warm area, called close-in warm area (CWA). In some cases, another so called distant warm area (DWA) might be found more downwind. [7]. Figure 2.3 shows some examples of the enhanced-U/V as it appears at the satellite imagery after application of the color enhancement.

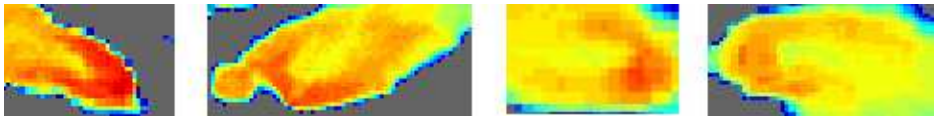


Figure 2.3: Examples of the enhanced-U/V cases, taken from the MSG satellite - IR channel, colder values are rendered red and orange, warmer values are rendered yellow and blue

There are several hypothesis about the enhanced-U/V's origin.[2, 7, 9, 8, 15]

The enhanced-U/V develops when a strong updraft penetrates into the lower stratosphere and results in an overshooting thunderstorm top. This overshooting top acts to block strong upper level winds and forces the flow to divert around it.

As the flow is diverted around the overshooting top, the main hypothesis is that the flow erodes the updraft summit and carries cloud debris downwind which is reflected in the colder equivalent blackbody temperatures (BT) of the enhanced-U/V feature.

The coldest BT, which is near the apex of the enhanced-V, is associated with adiabatic expansion due to the ascent of air parcels in the thunderstorm updraft region overshooting the tropopause.

Several hypotheses have been proposed to explain the warm region of BTs enclosed by the V-feature. [8, 9, 18, 21]

One hypothesis argues that the region is a result of subsidence of negatively buoyant overshooting cloud air downstream of an ascending cloud top.

A second hypothesis has been proposed which explains the warm region on the basis of radiative properties of the cloud particles. Based on radiative transfer simulations

and assuming that the ice water content varied spatially across the anvil, Heymsfield et al. found that the interior warm region had lower ice water content compared to the V-arms. This situation implies a smaller optical depth in the warm region and warmer BTs characteristic of lower altitudes.

Another theory argues that stratospheric cirrus generated in the wake of overshooting tops is sinking into the anvil. Located above an anvil top and at a warmer environmental temperature, the stratospheric cirrus appears warmer in the BTs sensed by the IR satellite channel.

The coldest point of the enhanced-U/V structure and the warmest point enclosed inside of the enhanced-V form a "cold-warm couplet". [5] There are several studies claiming possible link between the enhanced-U/V feature and supercells. Also the enhanced-U/V must be associated with a growing storm where growing stands for the enlargment of the colder IR contours or for lowering of the minimum cloud top temperature.[15] The last rule requires information about the storm cloud retrieved from a series of satellite images.

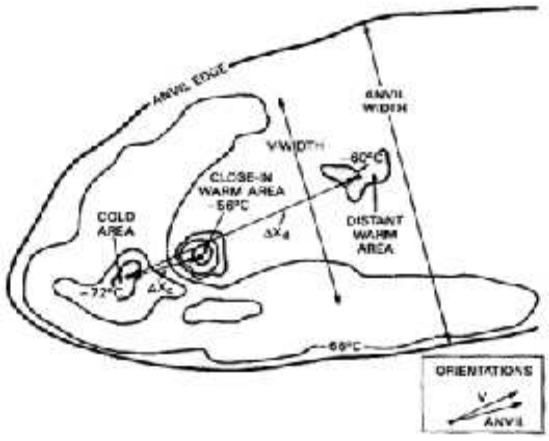


Figure 2.4: Scheme of the enhanced-U/V feature view from above [7]

2.2 Physical properties of the enhanced-U/V feature

Several quantitative analyses of the enhanced-U/V have been carried out(see [5, 15]). Brunner [5] has summarized and newly proposed several attributes which are bound to the enhanced-U/V occurrences. The parameters are TMIN - the minimum cloud top equivalent blackbody temperature (BT) observed from the satellite, TMAX - the maximum cloud top BT detected within an embedded warm area downwind of TMIN and TDIFF - the difference in cloud top BTs between TMIN and TMAX which forms the cold-warm couplet. TMIN median value was 200 K, TMAX median value was 217 K and TDIFF median value was 15 K. Heymsfield and Blackmer observed ranges of temperature difference between the cold and warm members of the cold-warm couplet

from 7 to 17 K [7, 5]. Orientation of the enhanced-U/V is influenced by the major direction of air flow and is mostly northern above Europe and above the USA. A median persistence of the feature is approximately 60 minutes. [15] Heymsfield and Blackmer [7] noticed that the coldest region associated with the cloud top is displaced upwind of the enhanced-U/V summit.

2.3 Satellite data description

This study is based on data obtained from the MSG (Meteosat 8 and 9) satellites, operated by the EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites). These satellites are geostationary, such satellites orbit above one specific point on the Earth's surface all the time and therefore their turnaround time is identical with the Earth's rotation. In the case of Meteosat 8, the point is placed at the Equator at 3.4 W, while Meteosat 9 is located at zero longitude. The height above Earth surface is 35,786 km and orbit radius is 42,155 km. Algorithm for enhanced-U/V detection processes only data from one channel - IR 10.8 which contains information about wavelengths between 9.80 and 11.80μm. Therefore intensity value of each pixel describes radiance of the area covered by the pixel. Data from MSG are received every 15 minutes. Satellite radiometer scans Earth disc from south towards north. One pixel of the image covers area of approximately 4x6km (at central Europe).

There are several factors which influence data image interpretation and processing. One of them is the fact that locations nearby disc edge are distorted by the **geostationary projection**. If the Earth is considered to be an oblate rotational ellipsoid than an inverse projection can be derived but an interpolation would have to be applied to guarantee image continuity. Detail information about the geostationary and inverse geostationary projection is described in [20].
Another feature which can cause misinterpretations is the **parallax**, which is an optical illusion which occurs in analogue meters and causes reading errors. It occurs when the viewing eye is not in the same plane, perpendicular to the meter face, as the indicating needle (see Figure 2.5). Therefore correction of the parallax error should be considered when specifying precise location of the storm's impact on the Earth's surface.
There is one more influence which might be included in the deliberation of the automatic enhanced-U/V detection. **Tropopause height** above Earth's surface depends on the latitude (see Figure 2.6), heightening of the tropopause 1 km reduces the tropopause temperature by 10K. Therefore, storm clouds in areas closer to equator can be significantly colder than those located far from equator.

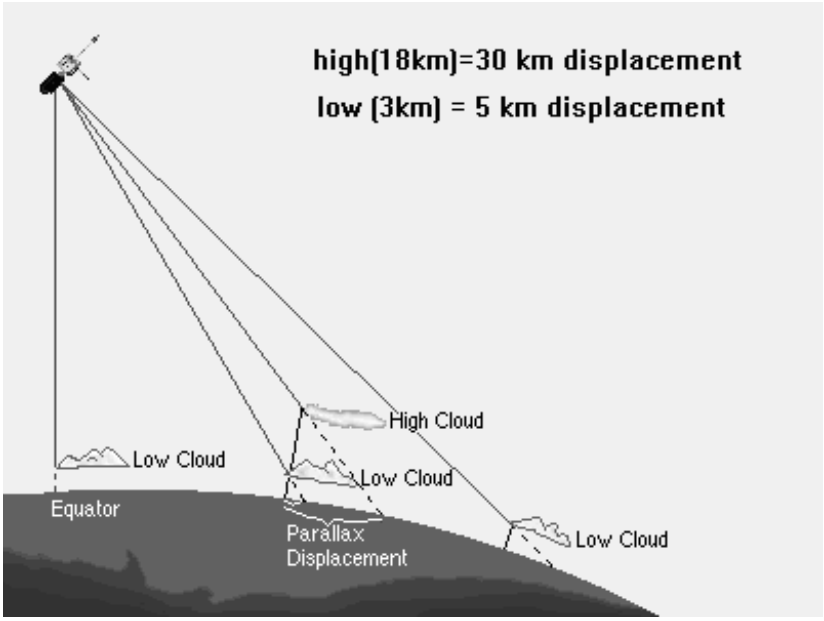


Figure 2.5: Scheme of parallax, [22]

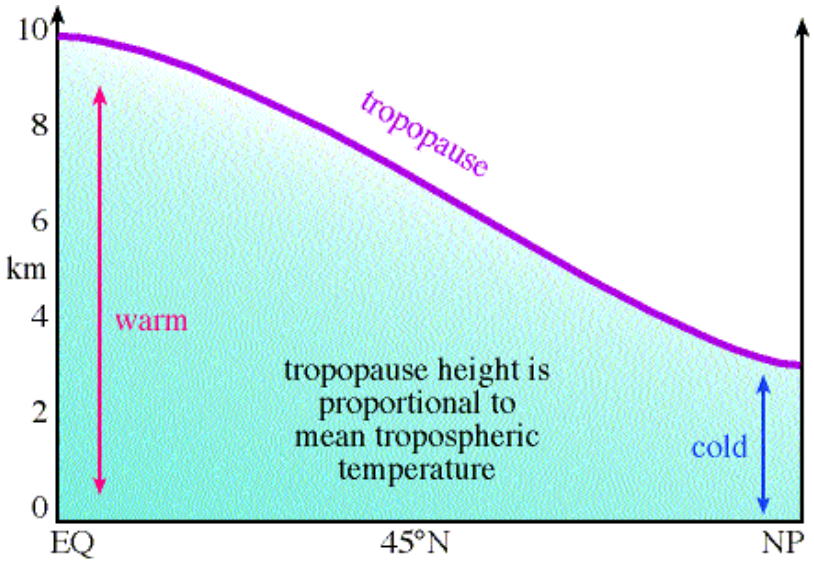


Figure 2.6: Scheme of change of tropopause height according to the latitude, [3]

Chapter 3

Algorithm proposal

This chapter proposes an approach to the automated enhanced-U/V detection. The first section aims to transform the meteorological motivation to the problem of the image processing, while the second section explains particular methods used by the algorithm for the enhanced-U/V detection, which are implemented in the prototype software.
The detection algorithm consists of two parts - a segmentation and a simple decision tree. Both the segmentation and the decision tree operate on various parameters extracted from enhanced-U/V candidates. Therefore, a major part of the proposal deals with the proper selection and extraction of these parameters.

3.1 Physical parameters of the enhanced-U/V candidates

Some of the parameters used by the algorithm naturally flow from the meteorological definition of the enhanced-U/V feature (see Section 2.2). Since each pixel of the image data encodes a temperature value, we can filter input image assuming that pixels whose value is bellow measured mean TMIN and above measured mean TMAX cannot be part of the sought feature.
Another physical property is the TDIFF value, which can vary between 7 K and 17 K. The computation of the TDIFF parameter for a concrete cloud depends on the parameters TMIN and TMAX measured for this concrete cloud. The TMAX value can be defined as the highest pixel value from the set of pixels "enclosed" by the candidate cloud, whereas the TMIN value can be obtained as the minimum pixel value "nearby" the summit of the candidate cloud. The details of the TMIN and TMAX seeking procedures is specified in 3.2.2.3.When the values of TMIN and TMAX parameters are evaluated, a simple difference between the TMAX and TMIN values determines the TDIFF parameter.
For the determination of the orientation parameter (as described in 2.2), a definition of the enhanced-U/V's axis is essential. The method for finding the axis is described in detail in 3.2.2.2. The orientation corresponds to the angle between the found axis and the x axis.

3.2 Algorithm structure

Satellite imagery only a simplified view on such a complex and hardly predictable feature as is the cloud development. Working with such a hardly definable problem brings significant uncertainty to the interpretation of results. The algorithm for the enhanced-U/V detection is designed with respect to the character of input data and aims to reduce the uncertainty to minimum. The priority of the proposal is to reduce the risk of omitting an enhanced-U/V occurrence whereas the false alarm elimination is solved with lower emphasis. Figure 3.2 demonstrates an example of a raw input image. The definition of the enhanced-U/V feature depends mainly on the local conditions, rather than on global information obtained from the image as a whole. The structure and properties of the storm clouds can therefore be studied independently for each cloud. This initial phase of separation is explained in the Section 3.2.1. Products of the isolation represent such parts of the storm clouds which have a chance to be the cold arms of an enhanced-U/V occurrence. Each such product must meet additional requirements to be stated as an enhanced-U/V candidate. The imposed requirements ensue both from the meteorological definition and from the shape demands. Many of those conditions can be expressed in the form of candidate's parameters and once the values of the parameters are assessed, the classification process, which decides whether the cloud is an appropriate candidate or not, can begin. Extraction of the features is described in the Section 3.2.2.
One of the possible and commonly used approaches to the classification is a simple decision tree. The choice of this approach was motivated by a small size of the training data set and the desired "true/false" decision about the candidates. Another possible method is based on probability theory and Section 5.2 provides further discussion about the approach.
The **decision tree** is a binary tree where every nonterminal node represents a decision. Depending upon the decision taken at such a node, control passes to the left or right subtree of the node. A leaf node then represents the outcome of taking the sequence of decisions given by the nodes on the path from the root to the leaf. [1]
For the purpose of the enhanced-U/V detection, a decision tree is constructed where each decision depends on a value of one selected parameter and whenever a condition is not fulfilled, the tested object cannot be a valid candidate. Figure 3.1 illustrates the decision process.

3.2.1 Segmentation

Segmentation is an operation which decides whether each one pixel of the input image belongs to an object of interest or not. Final product of such operation is a binary image.[11]
In the case of the enhanced-U/V detection, the product of segmentation should be the part of a cloud whose temperature corresponds to the predicted temperature of cold arms of the enhanced-U/V. Therefore the purpose of the segmentation is to remove the pixels which cannot belong to a storm cloud, which is realized by the step described in 3.2.1.1. Next method aims to remove the pixels which may form the warm part of the enhanced-U/V feature. This requirement is implemented by the actions described

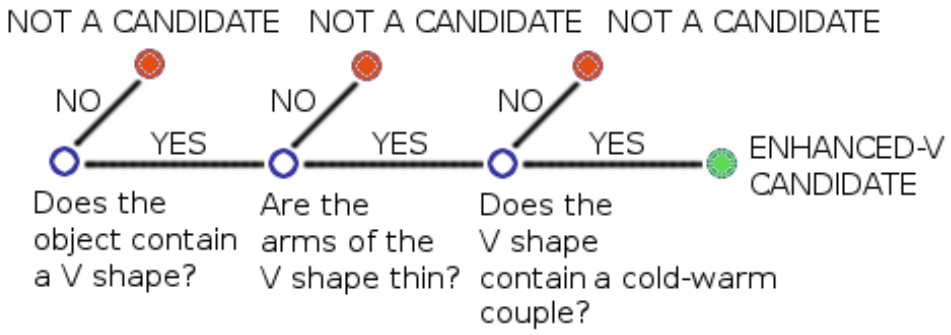


Figure 3.1: Decision tree used by the enhanced-U/V detection algorithm.

in 3.2.1.2.
The previous two actions tend to disrupt the original shape of the cold parts. Therefore two additional operations whose aim is to reconstruct the original shape are applied. The first of them is an operation called closing which is supposed to make objects more compact and to overlay artificially developed gaps. The second action is called median filter and is expected to make thborderer of objects smoother. The Sections 3.2.1.3 and 3.2.1.4 provide detail definition of the Closing and Median filter operations. Figure 3.6 illustrates the effect of different combinations of the Closing and Median filter operations.

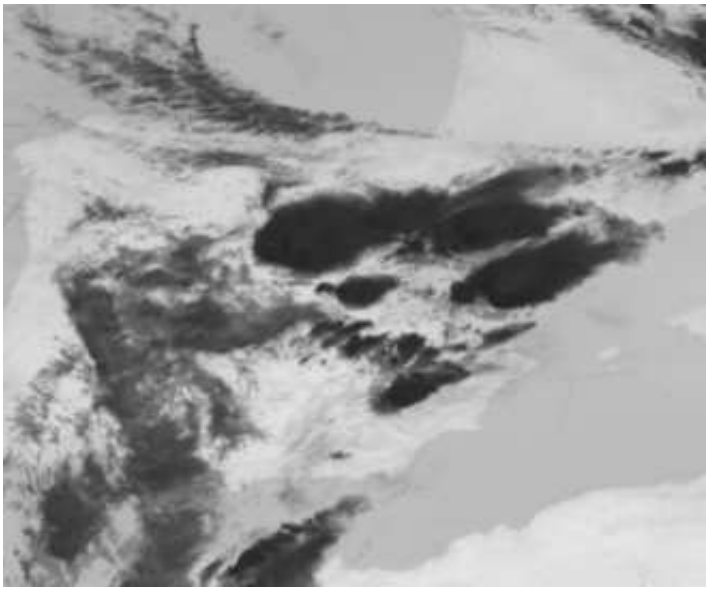


Figure 3.2: Part of the sample input image. Original values from the range of 0-65535 have been mapped to the range of 0-255 for purposes of presentation.

3.2.1.1 Basic thresholding

The point operation which marks a pixel to be the part of a sought object if its value lies inside of a selected interval (or lies above or bellow a selected value) and marks the pixel to be part of the background otherwise performs a simple segmentation technique is called **thresholding**. [12]
Quantitative analysis of the enhanced-U/V feature have shown that median value of the TMIN parameter is approximately 200K and median value of the TMAX parameter is approximately 220K (see Section 2.2 for details).
This interval might be a natural range for pixel values belonging to the storm clouds which can contain enhanced-U/V feature, but since the quantitative analysis measured only the storm occurrences over the middle latitudes and storm clouds get colder towards the Equator (see 2.6), the range of acceptable temperatures is enwiden to the interval 185-220K. This change is supposed to roughly eliminate influence of the tropopause height change (the height change is proportional to the change of the storm cloud temperature) nearby the Equator (see 2.3).
The discussion above implies to use a simple global thresholding with respect to the interval 185-220K.
A sample image processed by the threshloding method is shown in Figure 3.3.

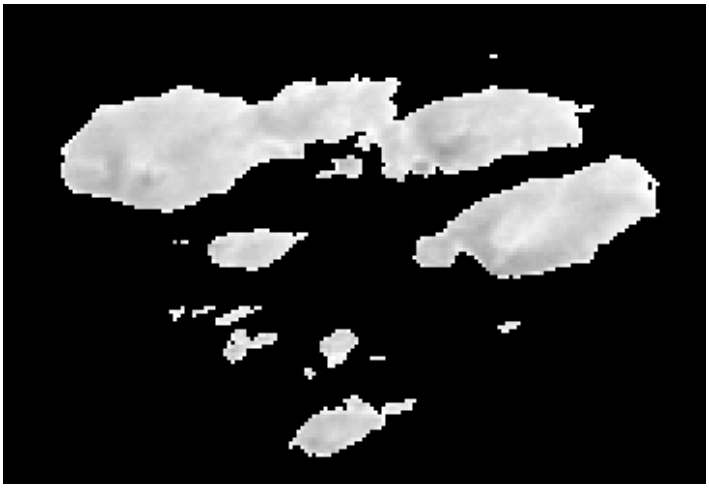


Figure 3.3: Basic thresholding result

3.2.1.2 Thresholding based on mean value

The basic thresholding outputs an image where each separate object is assumed to be either a single storm cloud or a connected group of a few storm clouds having similar temperature profiles. Next step of the consecution is the extraction of the coldest parts of the storm clouds, which may have a chance to be the cold arms of an enhanced-U/V occurrence. One possible solution of this problem is based on a simple statistic processing of the temperature values occurring within the tested storm cloud. The following steps are taken fore each separate cloud. At first, the procedure locates an unprocessed

cloud in the image. Secondly, value of each pixel of the selected cloud is stored. Then the **mean value** is computed from the stored values using the following formula:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{M} f_i x_i$$

Where N is the count of the pixels belonging to the tested cloud, M is the number of different temperature values (classes) found in the cloud and f_i is the number of pixels which values was analogous to the i^{th} temperature class.

When the mean value is known, the method filters out the pixels of the tested cloud whose value is higher than the mean value. After the method is applied for all storm clouds, the output image contains only the colder part of each storm cloud. This method expects the cold arms of the enhanced-U/V to be left and the warmer enclosed area to be removed. This technique is an adaptive form of the thresholding; the threshold value is obtained for each cloud independently. Figure 3.4 displays the statistics used by the procedure for a sample cloud. The pixels whose value corresponds to the temperature are left in the output image, the others are removed. Result of this thresholding is shown in Figure 3.5.

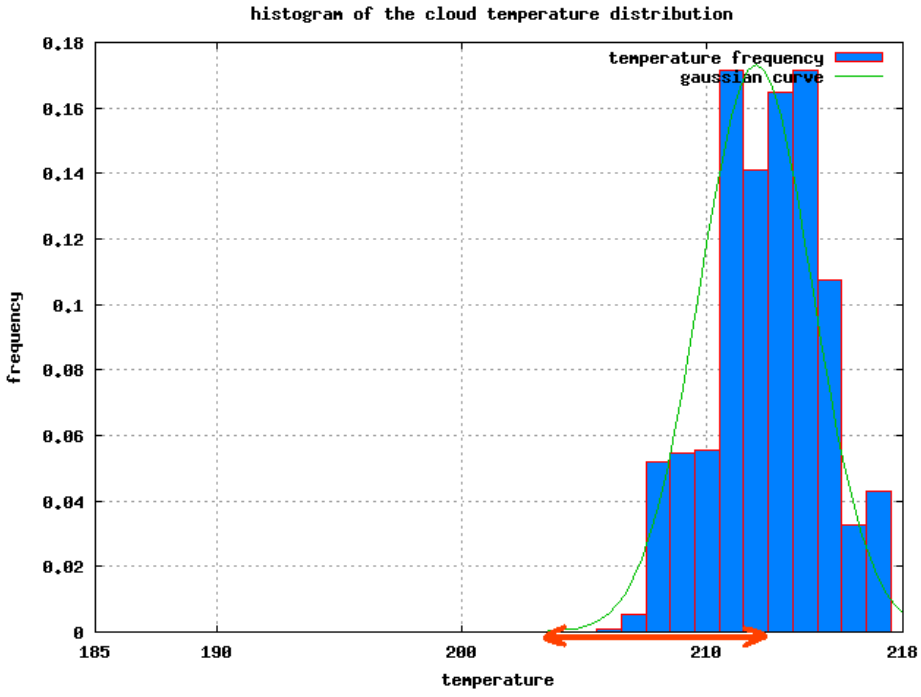


Figure 3.4: The histogram shows the full range of the temperature values which can be found within the sample cloud. Red arrow marks the interval of temperatures which are appropriate for further investigation.

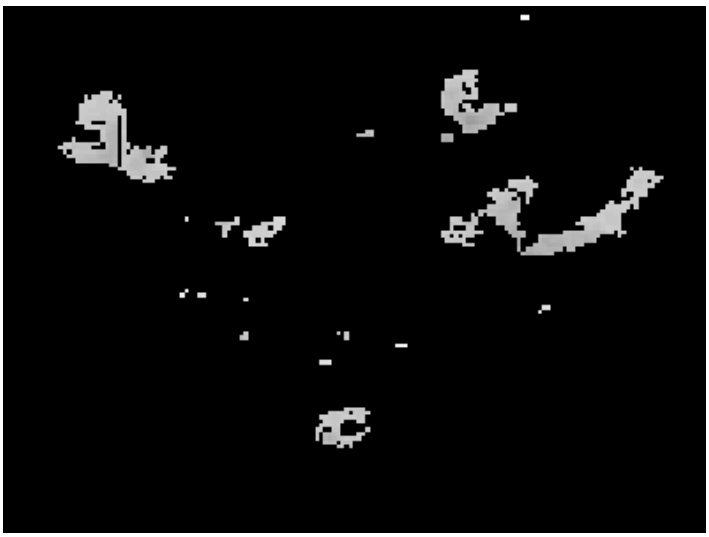


Figure 3.5: The figure demonstrates the result of thresholding based on median value when applied on the output image of the basic thresholding.

3.2.1.3 Closing

After the application of the advanced thresholding, the original shape of the cold cloud parts is often disrupted (see Figure 3.5). Therefore a method which would be able to reconstruct the shape and fill in the artificially created gaps is desired. A morphological operation called closing has characteristics similar to those in demand. Before this operation, the input image is converted to a binary image to simplify the following computation.
Closing consists of two simple operations - erosion and dilation. Both the erosion and the dilation use a matrix of values called the mask. G is regarded as the set of all the pixels of the image that are not zero. M is the set of the non-zero pixels in the mask. The symbol M_p denotes the mask shifted with its reference point (generally but not necessarily its center) to the pixel p.
Erosion is then defined as

$$G \ominus M = \{p : M_p \subseteq G\}$$

and **dilation** as

$$G \oplus M = \{p : M_p \cap G \neq \emptyset\}$$

.

The erosion of the set of pixels G by the set of pixels M can be expressed as the set of all pixels p for which M_p is completely contained in G . In contrast, the dilation of G by M is the set of all pixels for which the intersection between G and M_p is not an empty set. [13]

The dilation operator enlarges objects and closes small holes and cracks. Enlargement of objects by the size of the structure element can be reversed by a following erosion. This combination of operations is called the **closing** operation.

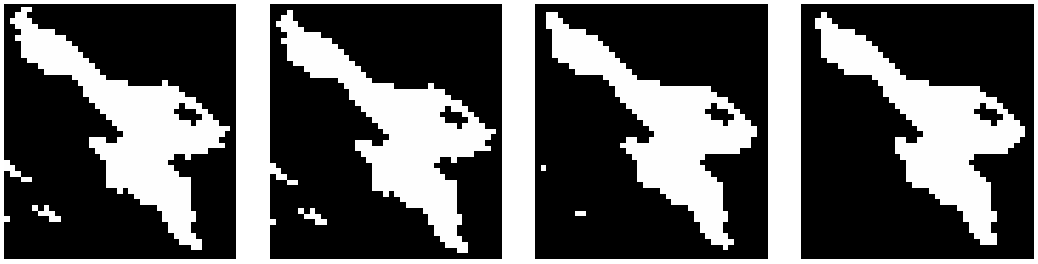
Closing is idempotent operation: $G \bullet M = (G \bullet M) \bullet M$ [13], which means that the repeated application of closing with the same mask doesn't make any further effect on the image (compare 3.6 (e) and (f)).

The mask used by the closing operation corresponds to the matrix :

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

This mask is placed on the image by its center point. Figures 3.6 (a) and (e) display the effect of the closing operation - objects become more compact.

(a) Input image (b) 1x Median f. (c) 2x Median f. (d) 3x Median f.



(e) 1x Closing (f) 2x Closing (g) 1x Closing ->1x Median f. (h) 1x Closing -> 2x Median f.

Figure 3.6: Effects of different combinations of Closing and Median filter to a sample object. The first image shows object before either closing or median filter. Numbers under the other images correspond to the number of iterations of Median filter and Closing operations. The combination of one application of Closing followed by one application of Median filter in our case provides the best results for the original shape reconstruction.

3.2.1.4 Median filter

The next step of the algorithm, which is applied right after the closing operation, aims to reduce the roughness of the object's border. One of the relevant methods, whose application makes an image smoother, is the median filter.

Median filtering is a nonlinear signal processing technique useful for noise suppression. Median filtering is performed by letting a mask move over the points of an image and replacing the value at the mask center with the median of the original values within the mask. This yields an output image which is usually smoother than the original one. The classical smoothing procedure is to use a linear low-pass filter and in many cases this is the most appropriate procedure. However, an important advantage of the median filter application is the fact that the operation preserves sharp edges, whereas linear low-pass filtering blurs edges. [14]

In the case of the enhanced-U/V detection algorithm, the median filter is applied to binary images. The results of the median filter application can be compared in Figure 3.6 (e) and (g). Repeated application of the median filter would bring better smoothness at the expanse of deformation of the object. The fact that the algorithm applies the median filter just once is substantiated by the argument that following parts of the algorithm use other methods for local border smoothing and thus the precise smoothing via the repeated median filtering is not necessary (see Figure 3.6 (g) and (h)).

3.2.2 Parameters extraction

Result of the segmentation part provides separate objects, each of which can contain cold arms of one or more enhanced-U/V features. For the purpose of identification of promissing enhanced-U/V candidates, a set of parameters of these objects needs to be defined, and methods used for the extraction of such parameters need to be developed. The term "enhanced-U/V" is derived from the shape of the sought feature. The shape is similar to the letter U or V. This characterization of the feature is rather subjective. However, some objective attributes which would be able to determine a V(U)-shaped object can be introduced. Two different approaches have been considered.
The first approach searches for a simple representation of an object called **skeleton**. The skeleton can be used for an easier interpretation of the shape of the object. Details of the skeleton definition can be found in [16]. Figure 3.8 shows samples of the skeleton structure. Although the skeleton provides a simplified view on the shape of the object, it is still a complex structure and recognition of a V-shaped part has shown to be rather difficult. Therefore another solution has been proposed.

The other approach analyzes the border of the object and tries to find **hollows**. The term hollow denotes a continuous area which belongs to the convex hull of the object but doesn't belong to the object itself. The area must also touch border of the convex hull (areas completely enclosed by the object are not considered to be hollows). The definition of the hollow is illustrated in Figure 3.7. When a hollow is marked, average width of the arm-shaped cloud parts which enclose the hollow is computed. The combination of the atributes describing shape of the hollow and the average arm width determines similarity of the candidate to the letter V. The method of extraction of those two parameters is described in detail in Section 3.2.2.1 and 3.2.2.2.
Another type of parameters reflects the physical properties of the enhanced-U/V feature. For each candidate hollow part of an object, the TDIFF parameter (see Sec-

tion 3.1) of the cold-warm couple can be estimated and orientation of the V shape measured. Sections 3.2.2.3 and 3.2.2.3 discuss methods used for computation of these parameters.

Figure 3.7: Illustration of the definition of the term hollow. Both green and red areas are not parts of the object but the red area is completly enclosed by the object and thus it is not considered to be a hollow.

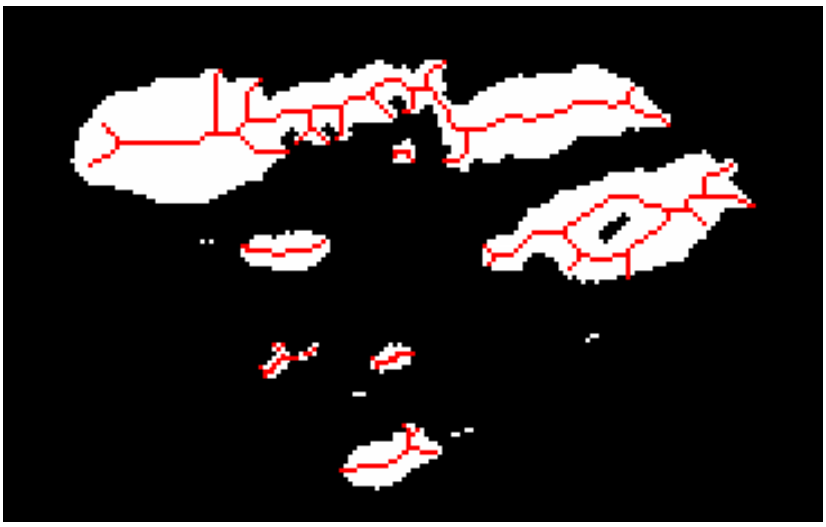


Figure 3.8: Sample objects and their skeletons.

3.2.2.1 Search for hollows

This section introduces a method for searching of V-shaped hollows in the object. A parameter which describes shape of a hollow is used later by the decision tree to select enhanced-U/V candidates form the set of all clouds (see the first node of the decision tree in Figure 3.1). The method of the parameter obtaining is based on the idea that each such hollow should consist of three parts - spike, pit and spike (marked as 1,2,3, resp. in Figure 3.9). The method is based on an analysis of convesity / concavity of the object's border. This approach is described by Horáček et al. in [10].

The algorithm uses a technique which expresses an approximate concavity at each point of the border. For the currently processed point of the border, the number of pixels lying in the intersection between the object and a circle mask placed by its center on the point is counted. The procedure is not supposed to consider the width of the object, its goal is to investigate the behaviour of the border curve only. Therefore, when measuring the intersection of the circle with an object, the part of the circle, which lies out of the object only because of the insufficient width of the object, is included in the intersection. The concavity value is then equal to the ratio of the intersection part to the whole circle area. The concavity values thus range between 0 and 1. Figure 3.9 demonstrates obtaining of the concavity value for one point of the border. The result values are connected to form a curve which reflects changes of the concavity along the border. If the border is longer than 50 points, the averaging of three neighbouring point values is applied. The averaging suppresses local noise and the output curve reflects the global trend rather than local changes.
The diameter of the circle depends on the area of the object, the dependency is expressed by the equation: $diameter = \sqrt{area}$. The relation is derived from two basic requirements for the diameter - the function should be growing, but it should grow faster for the lower values of the area and vice versa. The requirement of the slower diameter growth is motivated by the observation that the clouds large in their area come up by the merge of a few smaller clouds and are mainly oblong rather than rounded. A circle mask with large diameter would produce too general information about the border of such large clouds. The minimum possible diameter of the circle is set to five pixels.

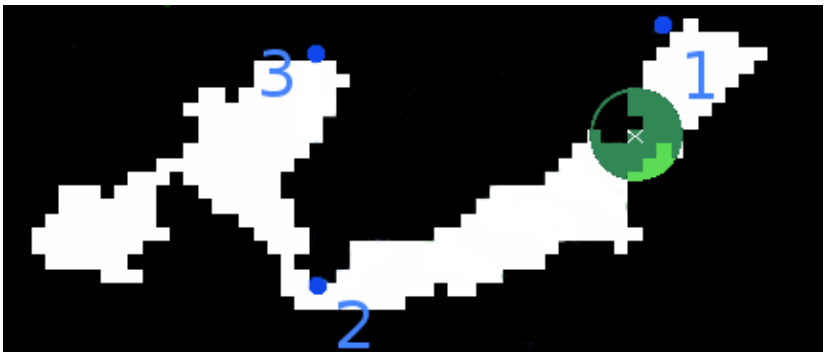


Figure 3.9: Illustration of the concavity measurement. The green area corresponds to the final intersection between the circle and object. Dark green represents the straight intersection with the object. Although the light green location lies outside of the object, it is included in the intersection because this situation is caused by the insufficient width of the object. Points 1, 2, 3 represents extremes of the concavity curve as implied by Figure 3.10.

The method then analyses the concavity curve and searches for the three consecutive segments of low concavity, high concavity and low concavity. These segments form a structure similar to a wave. The first and the third segments of the wave must

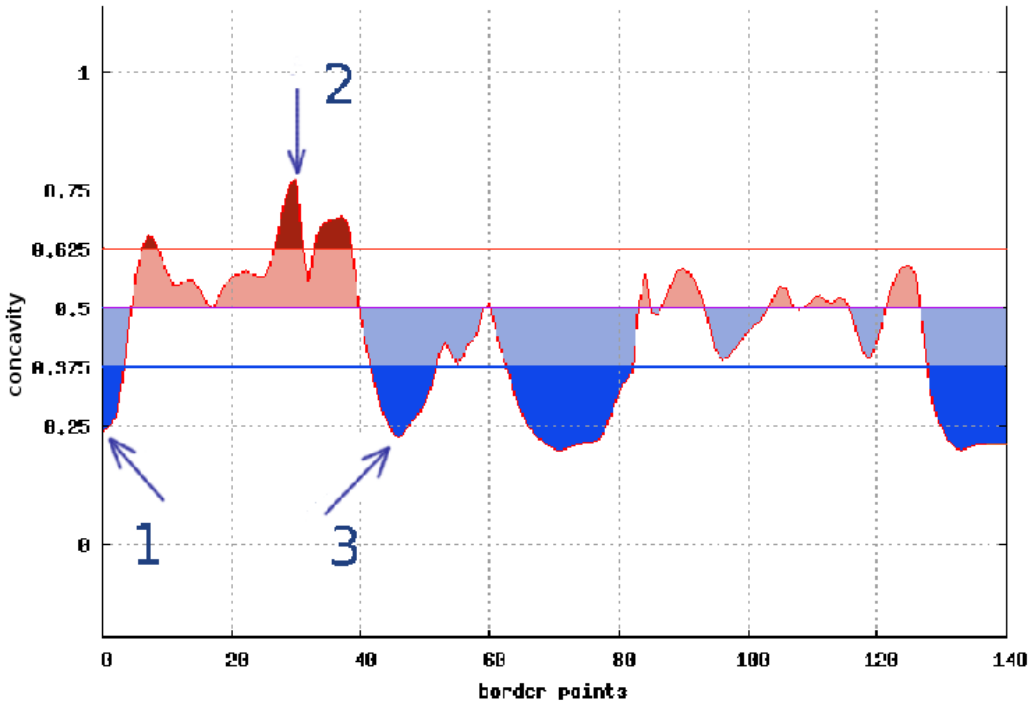


Figure 3.10: The concavity curve corresponding to the object shown in Figure 3.9. A wave encoding a candidate hollow must consist of three continued segments of low concavity - high concavity - low concavity, each of which must contain a point exceeding the threshold value. If translated to the colors of this figure, the wave must consist of a blue section(including a dark blue part) - a red section(including a dark red part) - a blue section(including a dark blue part)

reach the values of concavity lower than 0.375 whereas the second part of the wave must reach the value of concavity higher than 0.625. These thresholds have been set experimentally regarding to the training data set. The method also keeps information about the lowest point of the first and third segment and about the highest point of the second segment. Those extremes correspond to the points 1, 2, 3 in Figure 3.9. Because the curve basically starts at a random point of the border, the algorithm must consider the progression of the curve at the beginning when searching for the wave at the end and vice versa. Figure 3.10 displays a sample concavity curve. As a result, this method outputs hollows of the object which may represent candidate cold part of the enhanced-U/V. If there is no such candidate, the object is not included in further processing.

3.2.2.2 Determination of the average arm width

The output of the V-shaped hollow lookup provides a list of hollows which may represent the cold part of an enhanced-U/V. However, the lookup does not investigate the width of the cloud arms which form the hollows and thus a hollow may be similar to a deep indentation in an oval object as is illustrated in Figure 3.11 , but this type of

objects is not interesting for the enhanced-U/V detector. Therefore a procedure which can judge the average arm width for each candidate hollow is applied. This phase of the algorithm corresponds to the second step of the decision tree (see Figure 3.1).
The following steps are applied for each candidate hollow of the object. First, the axis of the hollow is identified. It is defined as a half line starting at the point 2 (the point with maximum concavity) and forms the axis of the angle $\measuredangle 123$.
Second, the average length of intersections of perpendiculars to the axis and the object provides information about the average arm width (see Figure 3.12). The length is measured in each border point which lies between border points 1,2 and 2,3. The lengths are averaged so the result represents the average arm width.
The last part of the procedure compares the measured value of the average arm width parameter to the depth of the hollow. If the depth is bigger than the average arm width, the hollow and the object continue to the next level of testing.
There are many possible definitions of the hollow depth. The one implemented in this study measures the distance of the point 2 from the line $\leftrightarrow 13$. This definition handicaps the V shapes whose one arm is longer than the other. Another types of depth measurement can behave differently, one of the other possible approaches defines depth as an average distance of the projection of points 1 and 3 on the axis of the hollow and point 2. Figure 3.13 presents several possible definitions of the hollow depth.
The described methods rejects candidates simillar to the object shown in Figure 3.11.

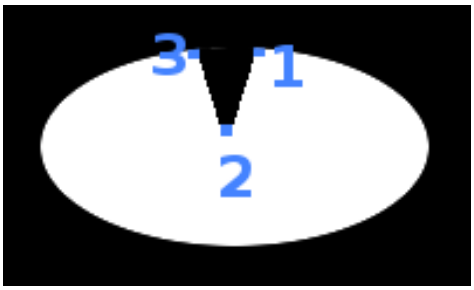


Figure 3.11: Example of a successful candidate of the V-shaped hollow lookup method. The method does not consider the width of the arms which form the hollow.

3.2.2.3 Seeking for the Cold-Warm couplet

The parameters metioned above are focused on the shape properties of the candidate hollows. Another parameter, which can be calculated from the candidate hollows, is based on the physical motivation and reflects the fact that the difference of temperatures of the coldest point of the cold part and the warmest point of the enclosed warm area must range between 7 K and 17 K (see Section 3.1). These two points are called the cold-warm couple. The comparison of the cold-warm couple temperature difference corresponds to the final step of the decision tree (see Figure 3.1).

The implementation of this condition uses both the original image and the output image of the segmentation. At first, the algorithm finds the points which fill the hollow

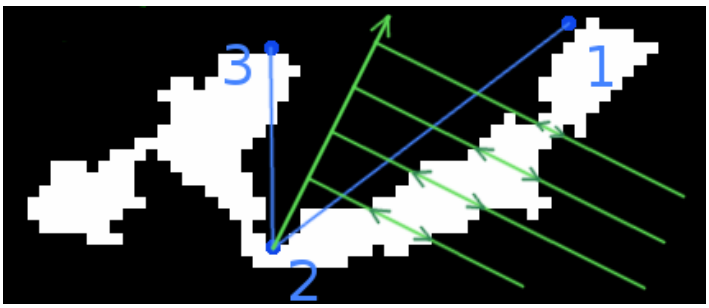
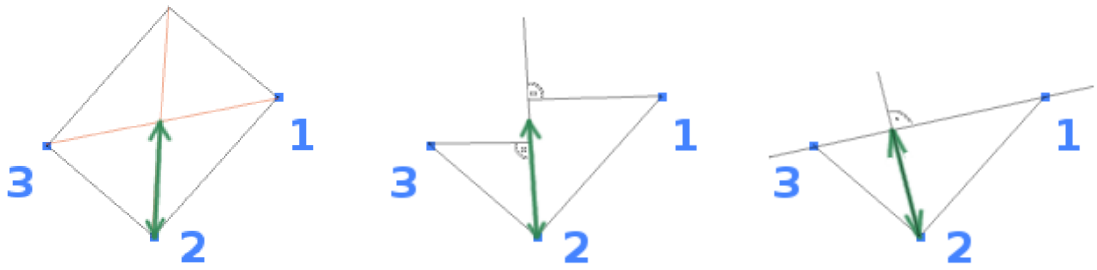


Figure 3.12: Demonstration of the average arm width measurement. The green lines represent axis of the hollow and perpendicular lines. Intersection of the perpendiculars and the object determines the average arm width.



(a) (b) (c)

Figure 3.13: Different ways of the hollow depth definition. Green arrow marks the depth. (a) The depth is defined as half length of the vector $\overrightarrow{d} = \overrightarrow{21} + \overrightarrow{23}$. This definition prefers hollows with unequally long arms. (b) The depth is the average distance of the projection of points 1 and 3 on the axis of the hollow and the point 2. (c) The depth is defined as the distance of the point 2 from the line $\leftrightarrow$ 13. This definition prefers hollows with similar arm length.

in the output image. The border of the hollow is defined as union of the border of the object between points 1,2 and 3 and the line 13. The method then localizes the area within the original image and intensities enclosed by this border are stored and sorted. Occurrences of temperature outliers which are not relevant for the enhaced-U/V detection is often a complication. Therefore, the chosen representative of the warm enclosed area is not the warmest point, but the point placed at the "90%" position (the point is the 90% quantile) of the sorted series of points of the enclosed area. The representative of the cold area is searched in a small surrounding of the hollow summit - corresponding to the point 2. This idea is based on the meteorological observations (see Section 3.1). A circle of the diameter equal to five points is placed on the point 2 and its intersection with the object specifies the area which contains the cold representative. The points inside of the intersection are sought in the original image and sorted by intensity. The representative of the cold area is the point placed at "10%" position (the point is the 10% quantile) of the sorted series of the points included in the intersection.

The value difference of these two representatives can range between 3 K and 20 K. The original temperature range 7-17 K has been extended so the influence of the different representative selection would be decreased. Figure 3.14 demonstrates the difference calculation. If the difference value meets the interval constraint, the object is stated as a successfull enhanced-U/V candidate.

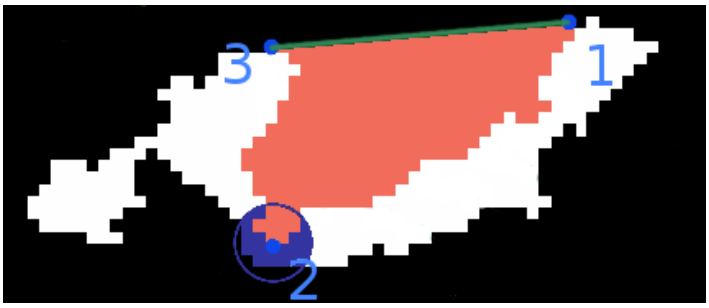


Figure 3.14: Illustration of the cold-warm couple lookup. The red area represents the enclosed area which contains the warm point. The enclosed area is bounded by the line $\leftrightarrow$ 13. The cold point is selected within the blue area.

3.2.2.4 Orientation calculation

Meteorological studies have implied another property of the enhanced-U/V which may be transformed to a parameter. It is the orientation of the candidate hollow. This orientation is influenced by the prevailing direction of the air flow, which is mostly constant at one geographical position e.g. the enhanced-U/V occurrences are often oriented towards the north above Europe and North America.
The orientation for each hollow can be expressed as the angle between the axis of the hollow and a line collateral with equator. Several tests were carried out to verify the orientation parameter applicability, but the results show that the orientation parameter can reject some of the enhanced-U/V occurrences marked by an expert. Therefore, the decision tree process does not use the orientation parameter. Though the parameter might be used by classification techniques different from the decision tree. The flow direction of the concrete storm cloud might also determine the parameter evaluation, but this improvement would request processing of the time series to calculate the direction of the storm cloud.

Chapter 4

Tests and experiments

An essential part of the algorithm development life cycle is the testing phase. The tests described in this chapter verify efficiency of the algorithm on a real data set, analyze inaccuracies of the detection and provide background for further improvement of the system.

The data used for testing were produced by an instrument called SEVIRI, which is a part of the MSG satellite system. The input images contain information about radiation measured at the area of northern 3/4 of the Northern Hemisphere (this area is from now on referred to as **the hemisphere**). The hemisphere includes a view on Europe, northern Africa, The Arabian Peninsula and the Atlantic. However, most of the tests were restricted to the European region. The first section (see 4.1) provides arguments for this constraint. The training data set consists of the data obtained during 28 spring/summer days of years 2005 and 2006. Data from each such day form a continuous series of images where the time difference between two consecutive images in the series is 15 minutes.

The training set of the enhanced-U/V occurrences includes only 10 cases, no other images were available during the development phase. Hence, the algorithm calibration based on such a small data set brings a high risk of the errors during the decisioning process. Another problematical aspect of the enhanced-U/V detector proposed in this study lies in the subjectivity of the enhanced-U/V definition. The subjectivity also causes inconsistency in the data evaluation provided by experts.

For the purpose of the clearer explanation of the test results, several terms are defined.
The term **Candidate** denotes a part of a storm cloud which was marked by the algorithm as a possible enhanced-UV.
A **Confirmed Candidate** is a Candidate which was marked as an enhanced-U/V instance by an expert too.
A **False Positive** is a Candidate which was rejected by an expert.
A **False Negative** is a part of a storm cloud which had been omitted by the algorithm but an expert marked it as the enhanced-U/V.
The term **Occurrence** denotes an enhanced-U/V instance in the real world (the storm

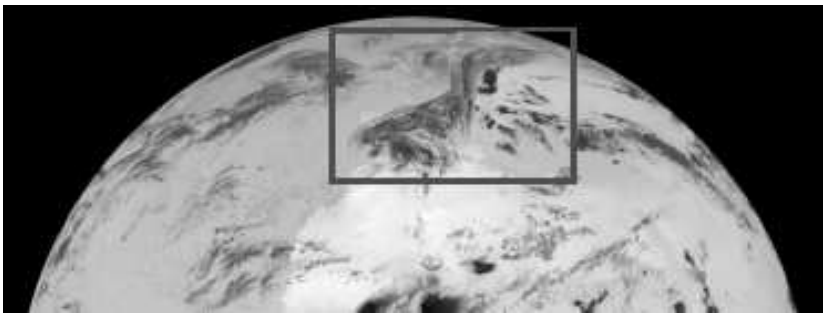


Figure 4.1: Illustration of the hemisphere. The gray rectangle denotes the European area.

cloud and its time development). An Occurrence can be observed in more than one image. Thus, if an enhanced-U/V occurence existed for 60 minutes, it could be found in four consecutive images.

The chapter consists of three sections. First of them introduces a test which provides arguments for restriction of the following experiments only to the European region. The next section summarizes and comments results of the enhanced-U/V detector application on a large testing data set. The last experiment investigates presence of an enhanced-U/V occurrence in a time series of images. This test investigates possible introduction of the enhanced-U/V time continuity as another classification parameter.

4.1 Test 1 - Observation area

The first test investigates the spatial arrangement of enhanced-U/V occurrences. The MSG satellite is placed nearby the zero meridian and above the Equator. Thus the distortion caused by the satellite projection is not that significant in the European area as it is in the areas located nearby the disk edge (e.g. Asia, the western part of the Atlantic). Besides that, the algorithm was calibrated on the training data set of the enhanced-U/V features observed above the European area. The temperature of the storm clouds is in average lower in the locations closer to the Equator than in Europe (see Section 2.3 for details).

This experiment compares images obtained during three randomly selected days (23.05.2005, 17.06.2006 and 12.07.2006). The data set consists of 119 shots displaying the hemisphere. This images were reviewed by an expert and the number of the candidates found in European region was compared to the number of those found in the rest of the hemisphere. The results are summarized in Table 4.1.

The table entries imply that most of the enhanced-U/V occurrences are located in the area of European continent whereas most of the candidates located in the rest of the hemisphere are false alarms. Moreover, only a few Confirmed Candidates occur outside Europe.

	Hemisphere	Europe	Hemisphere excl. Europe
Total # of Candidates	584	242	342
False Positives	450	112	338
Confirmed Candidates	134	130	4
Ratio of False Positives	77.05%	46.28%	98.83%

Table 4.1: Comparison of the observation results for the European area and the hemisphere. The column "Hemisphere" contains information about the whole hemisphere. The column "Europe" describes measured values for the European area only. The column "Hemisphere excl. Europe" denotes the area of the hemisphere excluding Europe. The first row provides the number of objects which were marked by the detection algorithm as possible enhanced-U/V candidates. The second and third rows of the table present results for the groups of candidates as defined in Chapter 4 introduction. The last row presents the ratio of the incorrectly marked objects to all marked objects.

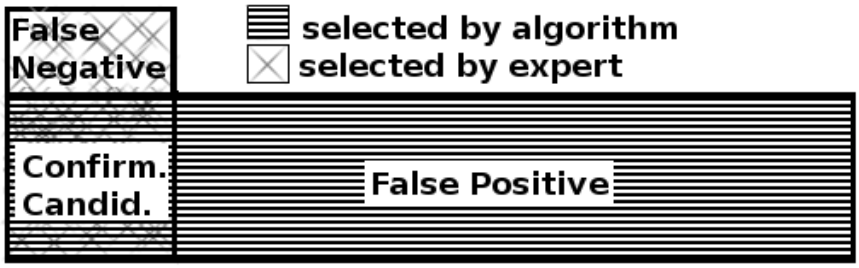


Figure 4.2: Illustration of the algorithm test results.

As a consequence, the following experiments process only the data obtained from the European area. Several techniques might be used for improvement of the algorithm behaviour in the areas outside Europe. The most obvious of them would compensate the projection distortion or reduce the influence of the atmospheric temperature structure (see Section 2.3 for detail description of these features).

4.2 Test 2 - Algorithm results

4.2.1 Test structure and results

This experiment is designed to test the accuracy of the enhanced-U/V detector. Data measured during 28 different spring/summer days form the testing data set; the data set contains 1113 images. With respect to the conclusions of the Section 4.1, the detector examined only the European region. The output images produced by the detector were examined by an expert. Figure 4.2 and Table 4.2 present results of the testing.

The results imply that approximately one third of the enhanced-U/V cases is omitted by the algorithm. Section 4.2.2 discusses possible causes of this result. A big portion of the candidates marked by the algorithm are false alarms, this detector behaviour is

Total # of Candidates	1398
False Positives	1153
False Negatives	142
Confirmed Candidates	245
Ratio of Confirmed Candidates to all enhanced-U/Vs	63%
Ratio of Confirmed Candidates to all Candidates	18%

Table 4.2: This table summarizes results of the detection accuracy test for the whole testing data set. Count of all objects which were marked by the detector as possible enhanced-U/V candidates is recorded in the row labeled "Total # of Candidates". The rows labeled "False Positives", "False Negatives" and "Confirmed Candidates" indicates number of objects corresponding to the groups as defined in Chapter 4 introduction. The row "Ratio of Confirmed Candidates to all enhanced-U/Vs" expresses the probability with which the detector marks an enhanced-U/V selected by an expert as a Candidate. This ratio is defined as (Confirmed Candidates)/(Confirmed Candidates+False Negatives). The last row "Ratio of Confirmed Candidates to all Candidates" contains information about the portion of the Candidates which corresponds to the Confirmed Candidates.

commented in the Section 4.2.3.

4.2.2 False negatives

Results of the test demonstrates that approximately one third of the enhanced-U/V occurrences were missed by the algorithm. This section aims to analyze possible causes of the False Negative cases. Each False Negative case must have been rejected at a certain algorithm stage. Once a candidate is rejected, no other stages of the algorithm are applied on it. Table 4.3 specifies the portion of the False Negatives rejected in particular algorithm phases.

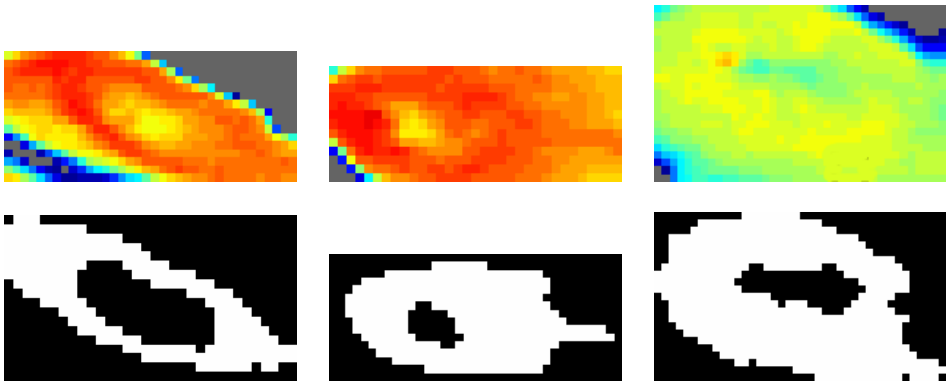


Figure 4.3: Samples of missed enhanced-U/Vs and their interpretation by the segmentation.

Although the initial segmentation phase seems to work correctly for the small training data set, it causes 41% cases of rejected data in the testing data set. Observation has

Training data		
Segmentation		100%
Classification	Concavity	0%
	Arm width	0%
	C-W couple	0%

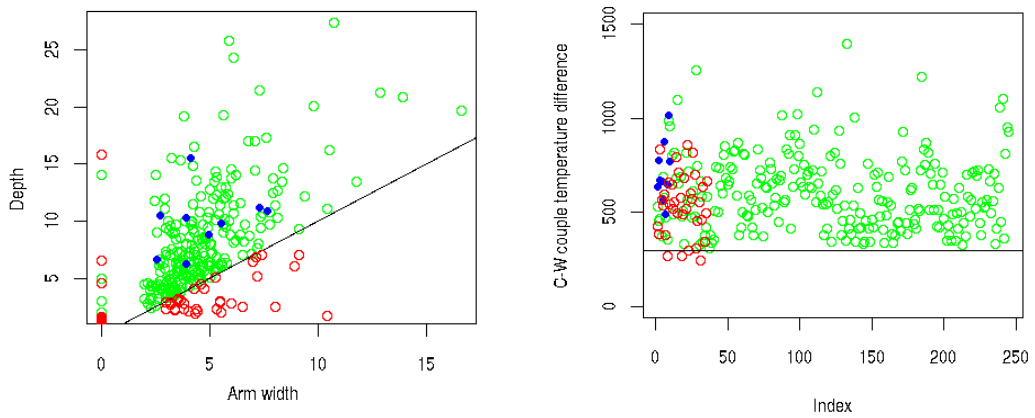
Testing data		
Segmentation		41%
Classification	Concavity	26%
	Arm width	23%
	C-W couple	10%

Table 4.3: The tables represent the distribution of the rejected cases in the algorithm phases for training and testing data sets. Each percentage specifies what portion of all False Negatives had passed all the previous steps but was rejected during the particular step. The "Segmentation" row specifies what portion of all False Negatives was either completely missed by the segmentation step (but marked by an expert as enhanced-U/V) or their shape was split in several pieces by the segmentation step. The "Concavity" row denotes a group of candidates whose rejection is caused by the unsuitable concavity curve shape (see Section 3.2.2.1). The "Arm width" row represents a group of candidates whose depth parameter reaches lower values than the average arm width parameter (see Section 3.2.2.2). The "C-W couple" corresponds to a group of objects whose cold-warm couple temperature difference is lower than 3K (see Section 3.2.2.3).

shown that many False Negatives developed during the advanced thresholding phase (see Section 3.2.1.1) by connection of the V-arms of the shape to a doughnut-like shape (see Figure 4.3). Another common distortion caused by the segmentation phase is the shortening of the V-arms. Such object often does not pass the classification criteria. Some of the objects are so shattered by the segmentation that they are not even examined by the classification phase. The segmentation extracts objects from the input image and determines their shape. This step of the algorithm influences the following classification phase, especially the concavity measuring step. This property makes the segmentation the key phase of the algorithm and its improvement may significantly improve the results of the classification phase too.

The classification phase may cause many False Negatives either. The phase consists of three main parts - concavity examination, comparison of the average arm width with depth and the cold-warm couple search (see Section 3.2.2). The False negatives caused by the concavity examination are mostly influenced by the segmentation phase as discussed in the previous paragraph. Influence of the other parts on the False Negatives development is demonstrated in Figure 4.4. Classification thresholds of those two parameters have been chosen according to the training data set, but particularly the rule based on comparison of the average arm width and depth seems to produce many False Negatives.

Two key factors have probably affected negatively the detection accuracy. First of them is a certain subjectivity of the enhanced-U/V definition which may lead to the differences in judgement of particular candidates either by different experts or by the same expert during different evaluation iterations. The second factor is a small size of the training data set which was available during the algorithm proposal and calibration.



(a) (b)

Figure 4.4: The figure illustrates rejection of objects during the classification phase. Blue dots represent the training set, green dots represent the Confirmed Candidates, and the red dots represent the False Negatives. The red dots lying above the threshold were rejected by another step of the algorithm. Figure (a) demonstrates the rule based on the average arm width, dependency between the average arm width and depth parameter is represented by the black line dividing the data set. Figure (b) the rule based on the temperature difference of the cold-warm couple. The threshold of 3K is represented by the horizontal line. Placement of the objects along the x axis does not provide any further information.

The detector is adjusted to search for objects similar by their attributes defined in the proposal to the objects in the training set.

4.2.3 False positives

The test has shown that many of the candidates marked by the detector do not fulfill all conditions required by an expert. A high number of the false alarms imply either possible incompleteness of the parameter set or wanting demands made by the current rules. However, the number of the False Negatives is too high even with the current settings and further restrictions would make the number even higher and reduction of the False Negatives has higher priority than reduction of the False Positives. Therefore, enrichment of the current parameter set by suitable parameters seem to be a better choice.

An important circumstance which is considered by experts but not by the algorithm is the development of the Occurrence in process of time. The algorithm proposal is supposed to examine only static images. Many structures which have properties similar to the enhanced-U/V features in a single image develop by connection or distortion of other storm clouds. An expert looks at the history of such object and rejects it. These False Positive cases can be detected only in a time series of images. Thus an algorithm adaptation which would provide information about a Candidate's development could

be used for a more accurate classification.

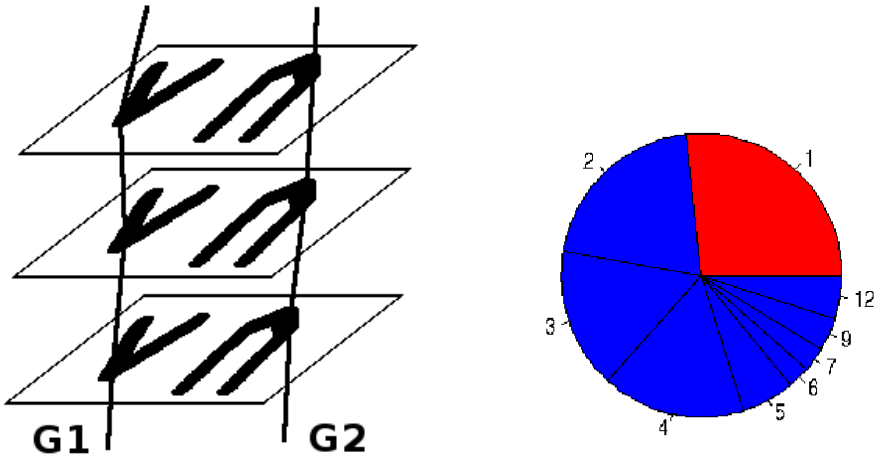
4.3 Test 3 - Time continuity of the observed enhanced-U/V features

This test aims to examine the time continuity of the Confirmed Candidates. This experiment is based on a meteorological observation which claims that the median persistence of the enhanced-U/V is 60 minutes (see Section 2.2). The frequency of the images used for the experiment is approximately one image per 15 minutes. Therefore, ideally, each occurrence of the enhanced-U/V would be marked in three or four consecutive images. The aim of this test is to verify that the algorithm also notes an enhanced-U/V occurence in more consecutive images. If the presumption of the time continuity is correct, the property could be used as another attribute which could reduce the number of False Positive cases.

However, the testing data are influenced both by the algorithm inaccuracy and the expert's subjectivity. Hence, a gap of two images (45 minutes) between two Confirmed Candidates corresponding to the same Occurrence (see definition in Section 4) was tolerated. The Confirmed Candidates were divided into groups by their pertinence to an Occurrence (Figure 4.5(a) illustrates the definition of the groups). Each group corresponds to one Occurrence and each member of a group must belong to a different image. Thus, a single member group represents a Confirmed Candidate for which no other Confirmed Candidate corresponding to the same Occurrence is found either in the three preceding images or in the three following images. A multimember group, on the other hand, represents a set of the Confirmed Candidates corresponding to the same Occurrence whose time distance to at least one another member of the group is less than three images. In other words, a single member group represents an Occurence which was identified by an expert (and by the algorithm) only in one image. A multimember group corresponds to an Occurence which was identified both by the expert and by the algorithm in more consecutive images.

The testing method was focused on the Confirmed Candidates only (objects marked both by the algorithm and an expert), the input data set contained 245 Confirmed Candidates. For each Confirmed Candidate three previus and three following images were examined manually and another Confirmed Candidate belonging to the same Occurence was sought in those six images. Figure 4.5 and Table 4.4 present results of the measurement.

The experiment has proved that a big portion of the Confirmed Candidates are members of a multimember group. This fact implies that the presence of another Candidate corresponding to the same storm cloud within the time interval $\pm$three images increases chances of a Candidate to become a Confirmed Candidate.



(a) (b)

Figure 4.5: (a)Illustration of the group definition. A group (an Occurrence) corresponds to a thread.(b) The figure demonstrates distribution of the correctly detected enhanced-U/V candidates in groups. The red part of the pie represents candidates which belong to a single member group. The blue part represents candidates belonging to a multimember group. The labels of each segment of the pie correspond to the total count of members of a group to which the candidate belongs.

Confirmed Candidates	245
# of groups	120
Average # of Confirmed Candidates in a group	2,04
Ratio of Confirmed Candidates belonging to a multimember group	72%

Table 4.4: This table records results of the time continuity test. The first row describes total number of the Confirmed Candidates. The second row indicates the number of the Occurrences. The third row represents the average number of the Confirmed Candidates per group. The last row of the table corresponds to the blue section in Figure 4.5.

Chapter 5

Future works

This study introduces a basic approach to the automatic enhanced-U/V detection. However, the development process brought many ideas for more improvements and extensions of the current algorithm. This chapter offers several topics which may be used as a groundwork of the future detector progression.

The first section is focused on the compensation of the satellite projection and influences of the atmospheric temperature structure. This step might widen the algorithm input from the European region to the whole hemisphere. The next section discusses possible replacement of the decision tree structure by a probabilistic approach. This change could decrease the number of the false negatives. Moreover, the output would provide wider possibilities for the judgement of the result than the decision tree provides now with the "true/false" candidate evaluation. Another section of this chapter is focused on the information obtainable from a time series. This information might provide a more complex view on the candidates selected from the static images. As a result, new parameters which could decrease occurrence of the false positives might be implemented. The last section discusses possible reduction of the expert estimation subjectivity. The aim of the section is to describe a suitable method for the decrease of the result dependency on a concrete person and to find the enhanced-U/V representatives which would be acceptable by most of the experts.

5.1 Preprocessing methods

During the testing phase, the area, which was examined for enhanced-U/V occurrences, was restricted from the hemisphere (the definition is introduced in Chapter 4) to the European area (see Section 4.1). This action was caused by different conditions for the enhanced-U/V detection in the center of the hemisphere and on the edges of the disc. This section discusses possible improvement of the algorithm behaviour in the problematic edge area.

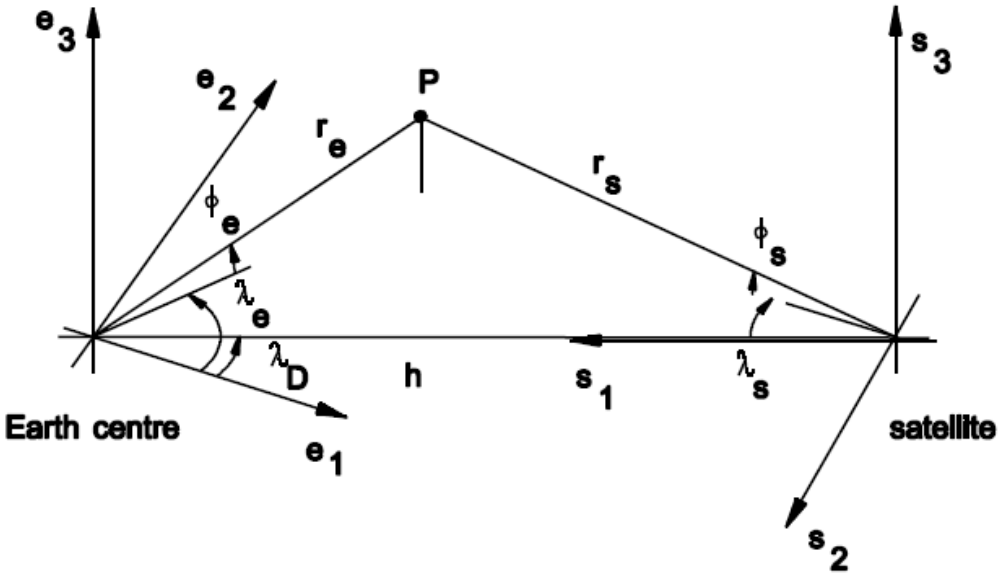


Figure 5.1: Illustration of the projection variables. Vector e_1 directs towards the zero meridian and touches the Equator. e_2 and e_3 complete the orthogonal system. P is a point on the Earth surface. Satellite coordinate system is defined by (s_1,s_2,s_3).[20]

5.1.1 Satellite projection

Images provided by a geostationary satellite actually represent projection of the Earth solid on a satellite plane. The torsion of the Earth surface causes distortion of objects in the locations nearby the disc edges. However, if several basic facts about the satellite position and projection are known, an inverse projection can be applied. The following procedure is assumed from the documentation of the input data format [20]. Figure 5.1 illustrates important variables used for the inverse projection.

The inverse transformation of a point $P = (x, y)$ of an input image to the longitude and latitude values (λ_e, ϕ_e) proceeds in the following steps:

1. Constants
h ... distance of the satellite from the Earth center
λ_d ... longitude of the satellite position
r_{eq} ... 6378.1690km
r_{pol} ... 6356.5853km
λ_s, ϕ_s ... satellite dependent values which can be calculated from the x and y values

2. Definition of the auxiliary Earth-related coordinate system (e_1', e_2', e_3')
The e_1' axis is defined by the Earth centre and the satellite centre, e_3 coincides with e_3 in Figure 5.1 and e_2' touches the Earth centre and is perpendicular to e_1' and e_3'. The Earth surface can be described using this Earth-related system as:

$$e'_1 = h - r_s cos(\phi_s) cos(\lambda_s) \tag{5.1}$$
$$e'_2 = -r_s cos(\phi_s) sin(\lambda_s) \tag{5.2}$$
$$e'_3 = r_s sin(\phi_s) \tag{5.3}$$

3. The Earth is described as an oblate rotational ellipsoid

$$\frac{e'^2_3}{r^2_{pol}} + \frac{e'^2_1 + e'^2_2}{r^2_{eq}} = 1 \tag{5.4}$$

4. Parameter r_s is expressed from the equations 5.1- 5.4

5. Expression of the longitude and latitude from (e_1', e_2', e_3')

$$\lambda_e = arctang\left(\frac{e'_2}{e'_1}\right) + \lambda_d \tag{5.5}$$
$$\phi_e = arctang\left(\frac{r^2_{eq}}{r^2_{pol}} \frac{e'_3}{\sqrt{e'^2_1 + e'^2_2}}\right) \tag{5.6}$$

Each point of the input image would be processed by the inverse projection. However, the output of the inverse projection contains gaps between the pixels. These gaps can be filled by an interpolation method. Basic methods of interpolation are described in [17].

The described method reduces distortions caused by the satellite projection and might improve the algorithm behaviour in the areas nearby disc edge.

5.1.2 Atmospheric temperature structure

This improvement is based on the physical properties of the atmosphere. Details are described in Section 2.3. The thickness of the troposphere atmospheric layer changes according to the changing latitude, its height reaches the greatest values nearby the Equator and the lowest values nearby poles. This structure influences an average temperature of the storm clouds so the storm clouds nearby the Equator are much colder than those located nearby poles.

This fact affects adversely the algorithm results because some of the detector phases use a temperature thresholding which is calibrated for the middle latitude bands. Therefore a temperature compensation of the parts lying outside the middle latitude bands might improve the algorithm results.

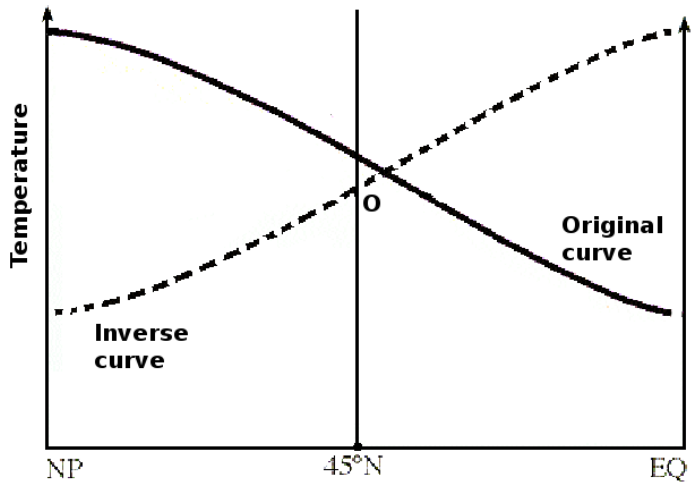


Figure 5.2: Illustration of the inverse temperature curve. The solid curve expresses the original dependency of the temperature on the latitude change. The inverse curve shifted to zero can be used for the temperature compensation. Point O denotes the new origin of coordinates after the shift.

The compensation could be done according to the curve which describes changes of the troposphere height with respect to the changes of latitude (see Figure 2.6). Since the temperature changes proportionally to the height, the curve can be transformed to express the change of temperature. The inverse curve could be shifted to the origin so the middle latitudes would not be affected by the compensation (see Figure 5.2). The compensation would change only the clouds. The result pixel of a cloud would be generated as a sum of the input pixel value + compensation value determined according to the pixels y position.

5.2 Classification

The classification part of the proposal is based on the decision tree method. The choice was motivated by a small size of the training set (10 enhanced-U/V cases). However the testing phase of this study prepared a significantly larger set of the enhanced-U/V samples (approximately 400 cases) which can be, in the future, used for the training purposes. Thus a more sophisticated classification method could be integrated into the algorithm proposal.

A classic approach, alternative to the decision tree, is based on the theory of probability. The parameter extraction part of the proposal would remain the same or could be enwiden by new parameters. The Bayesian classifier is a common representative of the probabilistic approach to the classification. A special case of the supervised learning with two known classes C_1 and C_2 is further assumed (the first class denotes the enhanced-U/Vs and the second denotes the objects which are not the enhanced-U/Vs).

The classifier then computes for both classes and for a new candidate described by the vector of measured features $\vec{x}$ the probability with which the new candidate belongs to the particular class. For a class C_i the probability can be expressed as:

$$P(C_i|\vec{x}) = \frac{p(\vec{x}|C_i).P(C_i)}{p(\vec{x}|C_1).P(C_1) + p(\vec{x}|C_2).P(C_2)} \tag{5.7}$$

Where $P(C_i)$ is called the prior probability. In the case of the detector, it can be computed as a ratio of the enhanced-U/V objects to all objects and vice versa. $p(\vec{x}|C_i)$ is than computed from the new training data set. Techniques used for the computation are described in [6].

As a result, the method determines the probability with which a new candidate object belongs to the class of enhanced-U/Vs. Such number provides much specific information about an object. The "true/false" result can be easily obtained by a simple thresholding e.g. an object would be marked by the algorithm as an enhanced-U/V candidate if the probability calculated by the classifier would reach 80%.

5.3 Time series

The current algorithm proposal is supposed to detect enhanced-U/Vs in a static image. However, when an expert evaluates the objects, the time development of the objects is considered. Thus, another logical step in the algorithm improvement is the investigation of the enhanced-U/V behaviour in a time series.

The acquisition frequency of the used data is approximately 15 minutes. The change of the storm cloud positions during this time is relatively small and thus a tracking method which would observe the cloud development could be introduced.

Some meteorological observations implied that a mean persistence of the enhanced-U/V is 60 minutes (see Section 2.2). Therefore, if the image acquisition frequency is less than 30 minutes, an enhanced-U/V may be found in more consecutive images. An experiment, which should examine the claimed time persistence of the enhanced-U/V features, was carried out (see Section 4.3). Candidates which were marked both by the algorithm and by an expert were tracked in a time series and other candidates selected both by the algorithm and by an expert which would belong to the same enhanced-U/V were sought. The test has shown that most of the enhanced-U/V occurrences were observed repeatedly. This fact might be used as another decision rule of the classification.

Another observation claims that an enhanced-U/V should be a part of a growing storm cloud ([15]). Whenever a tracking method is implemented, this property of a candidate's predecessor can be measured.

The orientation of an enhanced-U/V occurrence is determined by the direction of the storm cloud move. A vector of the movement could be derived from the time series

and consequently the orientation parameter could be used by the classification.

The methods described in this section could lead to a significant decrease of the False Positive ratio.

5.4 Subjectivity reduction

The certain subjectivity hidden in the enhanced-U/V definition causes complications both during the development and testing phase. Therefore, introduction of some methods which would reduce the subjectivity influence is desired.

In general, repeated evaluation and following averaging of the results might provide a more objective view on data. Hence, even the training data set should consist of the enhanced-U/Vs which were repeatedly marked by different experts. Such training data set would contain only the enhanced-U/V occurrences which are acceptable for all the experts. The basic algorithm calibration could be adapted to this "core" training set. This method may improve accuracy of the detector and thus the data set creation is highly desired.

Chapter 6

Conclusions

The analysis of the satellite imagery is a significant component of the weather forecasting. However, many of the analytic techniques are still human dependent. One of such techniques is the enhanced-U/V observation. This study investigated basic properties of the enhanced-U/V feature from the point of view of computer science and following automatic detection.

The main contribution of the thesis was the transformation of the meteorological problem to the problem of image processing and design of the intial algorithm structure. The pilot application created on the basis of the algorithm proposal allowed verification of the method suitability and accuracy. The results obtained during the testing phase provided valuable information about the algorithm behaviour and the findings can be used as a groundwork for further development.

In summary, the implemented algorithm is able to find approximately two thirds of the desired enhanced-U/V occurrences, but on the other hand the portion of all candidates which are false positives is very high. The final part of the study suggested several ways of reducing the number of both false postives and false negatives. The implementation of the improvements could markedly better the detection results.

As a whole, the study provided a basic resarch necessary for the successful ehanced-U/V automatic detection.

Appendix

Pilot application

An important part of this study is the pilot application which has been developed on the basis of the proposal.

The application was developed in the C++ language and its target system is Linux. The input format of the satellite images processed by the program is the "xpif" format. The format must contain information about the radiation in each pixel.

The experiments with the algorithm were carried out on the following PC configuration:
processor AMD Athlon(tm) XP, 1.2 GHz
750 MB RAM
An average time spent on an image with restriction to the European region was 10 seconds, a processing of the whole hemisphere took approximately 30 seconds.
The pilot application is included in this thesis.

Bibliography

[1] A dictionary of computing. Oxford Reference Online, 2004. http://www.oxfordreference.com/views/ENTRY.html?subview=Main&entry=t11.e1305.

[2] R. F. Adler, M. J. Markus, and D. D. Fenn. Detection of severe midwest thunderstorms using geosynchronous satellite data. *Monthly Weather Review*, 113:769–781, 1985.

[3] N. Atkins. http://apollo.lsc.vsc.edu/classes/met130/notes/chapter8/trop_hght.html.

[4] M. Branick. A comprehensive glossary of weather terms for storm spotters. NOAA Technical Memorandum NWS SR-145. http://www.srh.noaa.gov/oun/severewx/glossary.php.

[5] J. C. Brunner, S. A. Ackerman, A. S. Bachmeier, and R. M. Rabin. A quantitative analysis of the enhanced-v feature in relation to severe weather. *Weather and Forecasting*, 2007.

[6] R. O. Duda, P. E. Hart, and D. G Stork. *Pattern Classification*. John Wiley & Sons, 2nd ed. edition, 2001.

[7] G. M. Heymsfield and R. H. Blackmer, Jr. Satellite-observed characteristics of midwest severe thunderstorm anvils. *Monthly Weather Review*, 116:2200–2224, 1988.

[8] G. M. Heymsfield, R. H. Blackmer, Jr., and S. Schotz. Upper level structure of oklahoma tornadic storms on 2 may 1979, pt. 1 radar and satellite observations. *Journal of Atmospheric Science*, 40:1740–1755, 1983.

[9] G. M. Heymsfield, G. Szejwach, S. Schotz, and R. H.. Blackmer, Jr. Upper level structure of oklahoma tornadic storms on 2 may 1979, pt. 2 proposed explanation of "v" pattern and internal warm region in infrared observations. *Journal of Atmospheric Science*, 40:1756–1767, 1983.

[10] O. Horáček, J. Kamenický, and J. Flusser. Recognition of partially occluded and deformed binary objects. Proceedings of the 11th International Conference on Computer Analysis of Images and Patterns. CAIP 2005:415–422, 2005.

[11] B. Jähne. *Digital Image Processing*, volume IV, chapter 16, pages 449–462. Springer Berlin Heidelberg, 2005.

[12] B. Jähne. *Digital Image Processing*, volume II, chapter 10, pages 257–295. Springer Berlin Heidelberg, 2005.

[13] B. Jähne. *Digital Image Processing*, volume V, chapter 18, pages 501–514. Springer Berlin Heidelberg, 2005.

[14] B. I. Justusson. *ed. Two-dimensional Digital Signal Processing II: Transforms and Median Filters*, volume 43 of *Topics in Applied Physics*, chapter 5 Median Filtering: Statistical Properties, pages 161–196. Springer New York, 1981.

[15] D. W. McCann. The enhanced-v: A satellite observable severe storm signature. *Monthly Weather Review*, 1983.

[16] W.K Pratt. *Digital image processing : PIKS inside*, chapter 14. John Wiley & Sons, 3rd ed. edition, 2001.

[17] W.K Pratt. *Digital image processing : PIKS inside*, chapter 13. John Wiley & Sons, 3rd ed. edition, 2001.

[18] M. Setvak, R. M. Rabin, and P. K. Wang. Contribution of modis instrument to the observations of deep convective storms and stratospheric moisture detection in goes and msg imagery. *Atmospheric Research*, 2005.

[19] E. S. Takle. http://www.meteor.iastate.edu/gccourse/atmos/atmos_lecture.html.

[20] VCS Aktiengesellschaft. *Reference Manual 2met! XPIF, Extended Processed Image File Architecture*, issue 1.5 edition, March 2005.

[21] P. K. Wang, H-m. Lin, S. Natali, S. Bachmeier, and R. Rabin. Cloud model interpretation of the mechanisms responsible for the satellite-observed enhanced v and other features atop some midwest severe thunderstorms. *Proceedings of the 11th American Meteorological Society Conference on Cloud Physics*, 2002.

[22] J. Wick. http://en.wikipedia.org/wiki/Image:Parallax_Example.svg.